

B72

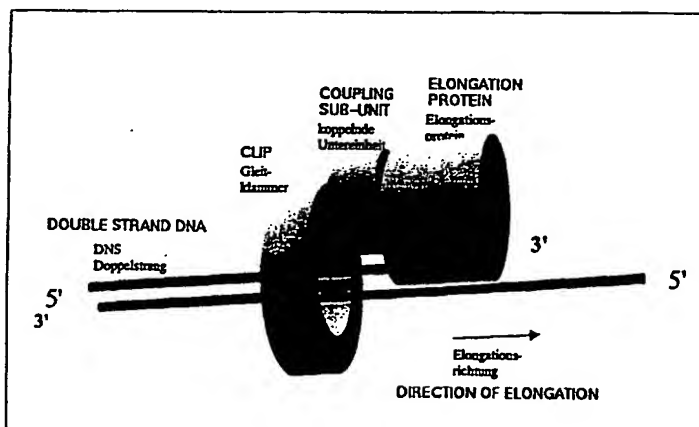


PCT
WELTORGANISATION FÜR GEISTIGES EIGENTUM
Internationales Büro
INTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE
INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

<p>(51) Internationale Patentklassifikation 7 : C12N 15/31, 9/12, C07K 14/195, C12P 19/34, C12Q 1/68</p>	A2	<p>(11) Internationale Veröffentlichungsnummer: WO 00/08164</p> <p>(43) Internationales Veröffentlichungsdatum: 17. Februar 2000 (17.02.00)</p>									
<p>(21) Internationales Aktenzeichen: PCT/DE99/02480</p> <p>(22) Internationales Anmeldedatum: 6. August 1999 (06.08.99)</p> <p>(30) Prioritätsdaten:</p> <table style="width: 100%;"> <tr> <td>198 35 653.6</td> <td>6. August 1998 (06.08.98)</td> <td>DE</td> </tr> <tr> <td>198 40 771.8</td> <td>7. September 1998 (07.09.98)</td> <td>DE</td> </tr> <tr> <td>99111795.3</td> <td>18. Juni 1999 (18.06.99)</td> <td>EP</td> </tr> </table> <p>(71) Anmelder (für alle Bestimmungsstaaten ausser US): LION BIOSCIENCE AG [DE/DE]; Im Neuenheimer Feld 517, D-69120 Heidelberg (DE).</p> <p>(72) Erfinder; und</p> <p>(75) Erfinder/Anmelder (nur für US): KILGER, Christian [DE/DE]; Handschuhheimer Landstrasse 47, D-69121 Heidelberg (DE). KOBER, Ingo [DE/DE]; Am Großen Wald 6, D-69251 Gaiberg (DE). VOSS, Hartmut [DE/DE]; Birkenweg 8A, D-69221 Dossenheim (DE). MOECKEL, Gerd [DE/DE]; Muehlenstrasse 20, D-68723 Oftersheim (DE).</p> <p>(74) Anwälte: GODDAR, Heinz usw.; Boehmert & Boehmert, Hollerallee 32, D-28209 Bremen (DE).</p>	198 35 653.6	6. August 1998 (06.08.98)	DE	198 40 771.8	7. September 1998 (07.09.98)	DE	99111795.3	18. Juni 1999 (18.06.99)	EP	<p>(81) Bestimmungsstaaten: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO Patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</p> <p>Veröffentlicht Ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts.</p>	
198 35 653.6	6. August 1998 (06.08.98)	DE									
198 40 771.8	7. September 1998 (07.09.98)	DE									
99111795.3	18. Juni 1999 (18.06.99)	EP									

(54) Title: THERMOSTABLE *IN VITRO* COMPLEX WITH POLYMERASE ACTIVITY

(54) Bezeichnung: THERMOSTABILER *IN VITRO*-KOMPLEX MIT POLYMERASEAKTIVITÄT



(57) Abstract

The inventive thermostable *in vitro* complex for template-dependent elongation of nucleic acids comprises a thermostable staple protein and a thermostable elongation protein.

(57) Zusammenfassung

Der erfindungsgemäße thermostabile *in vitro*-Komplex zur Template-abhängigen Elongation von Nukleinsäuren umfaßt ein thermostabiles Gleitklammerprotein und ein thermostabiles Elongationsprotein.

LEDIGLICH ZUR INFORMATION

Codes zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

AL	Albanien	ES	Spanien	LS	Lesotho	SI	Slowenien
AM	Armenien	FI	Finnland	LT	Litauen	SK	Slowakei
AT	Österreich	FR	Frankreich	LU	Luxemburg	SN	Senegal
AU	Australien	GA	Gabun	LV	Letland	SZ	Swasiland
AZ	Aserbaidschan	GB	Vereinigtes Königreich	MC	Monaco	TD	Tschad
BA	Bosnien-Herzegowina	GE	Georgien	MD	Republik Moldau	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagaskar	TJ	Tadschikistan
BE	Belgien	GN	Guinea	MK	Die ehemalige jugoslawische Republik Mazedonien	TM	Turkmenistan
BF	Burkina Faso	GR	Griechenland	ML	Mali	TR	Türkei
BG	Bulgarien	HU	Ungarn	MN	Mongolei	TT	Trinidad und Tobago
BJ	Benin	IE	Irland	MR	Mauretanien	UA	Ukraine
BR	Brasilien	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Island	MX	Mexiko	US	Vereinigte Staaten von Amerika
CA	Kanada	IT	Italien	NE	Niger	UZ	Usbekistan
CF	Zentralafrikanische Republik	JP	Japan	NL	Niederlande	VN	Vietnam
CG	Kongo	KE	Kenia	NO	Norwegen	YU	Jugoslawien
CH	Schweiz	KG	Kirgisistan	NZ	Neuseeland	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Demokratische Volksrepublik Korea	PL	Polen		
CM	Kamerun	KR	Republik Korea	PT	Portugal		
CN	China	KZ	Kasachstan	RO	Rumänien		
CU	Kuba	LC	St. Lucia	RU	Russische Föderation		
CZ	Tschechische Republik	LJ	Liechtenstein	SD	Sudan		
DE	Deutschland	LK	Sri Lanka	SE	Schweden		
DK	Dänemark	LR	Liberia	SG	Singapur		
EE	Estland						

Thermostabiler *in vitro* Komplex mit Polymeraseaktivität

Beschreibung

5

Die Erfindung betrifft einen thermostabilen *in vitro*-Komplex zur Template-abhängigen Elongation von Nukleinsäuren, einen thermostabilen *in vitro* Komplex sowie dafür kodierende DNA-Sequenzen und Vektoren. Die Erfindung betrifft weiterhin die Verwendung der erfindungsgemäßen Komplexe in

10 Verfahren zur Template-abhängigen Elongation von Nukleinsäuren, wie PCR Reaktionen reversen Transkription, DNA-Markierung oder DNA-Sequenzierung, bei denen *in vitro* Template-abhängige DNA Strangsynthese erfolgt. Schließlich betrifft die Erfindung noch Kits bzw. Reagenzienkits zur Durchführung der erfindungsgemäßen Verfahren.

15

DNA-Polymerasen gehören zu einer Gruppe von Enzymen, die einzelsträngige DNA als Template für die Synthese eines komplementären DNA Stranges verwenden. Diese Enzyme spielen eine bedeutende Rolle im Nukleinsäurestoffwechsel, einschließlich der Prozesse DNA-Replikation,

20 Reparatur und Rekombination. DNA-Polymerasen wurden in allen zellulären Organismen identifiziert, von bakteriellen bis zu menschlichen Zellen, in vielen Viren sowie in Bakteriophagen (Komberg, A. & Baker, T. A. (1991) DNA Replikation WH Freeman, New York, NY). Man faßt in der Regel die Archaeobakterien und die Eubakterien zusammen zu der Gruppe der

25 Prokaryonten, der Organismen ohne echten Zellkern, und stellt ihnen die Eukaryonten, die Organismen mit echtem Zellkern, gegenüber. Gemeinsam sind vielen Polymerasen aus den verschiedensten Organismen oftmals Ähnlichkeiten in der Aminosäuresequenz sowie Ähnlichkeiten in der Struktur (Wang, J., Sattar, A.K.M.A.; Wang, C.C., Karam, J.D., Konigsberg, W.H. &

30 Steitz, T.A. (1997) Crystal Structure of pol α family replication DNA

polymerase from bacteriophage RB69. Cell 89, 1087-1099). Organismen wie der Mensch besitzen eine Vielzahl von DNA-abhängigen Polymerasen, von denen jedoch nicht alle für die DNA-Replikation zuständig sind, sondern einige auch DNA-Reparatur durchführen. Replikative DNA-Polymerasen bestehen *in vivo* meist aus Proteinkomplexen mit mehreren Einheiten, welche die Chromosomen der zellulären Organismen und Viren replizieren. Eine generelle Eigenschaft dieser replizierenden Polymerasen ist im allgemeinen eine hohe Prozessivität, das heißt, deren Fähigkeit, Tausende von Nukleotiden zu polymerisieren ohne vom DNA Template abzudissoziieren (Kornberg, A. & Baker, T. A. (1991) DNA Replikation. WH Freeman, New York, NY).

Im Stand der Technik sind hochprozessive Replikationsmechanismen bekannt, wobei es sich dabei zum einen um zelluläre Mechanismen und zum anderen um die bei den Bakteriophagen T4 und T7 auftretenden Replikationsmechanismen handelt.

Der Replikationsapparat umfaßt eine Vielzahl von Komponenten. Dazu gehören, unter anderem, a) Polymeraseaktivität aufweisende Proteine, b) Proteine, die an der Ausbildung einer Klammerstruktur beteiligt sind, wobei der Klammerstruktur unter anderem die Aufgabe zukommt, eine Polymeraseaktivität an ihr Template zu binden, die Bindung zu stabilisieren und somit die Dissoziationskonstante entsprechend zu ändern, c) Proteine, welche die Klammer auf das template laden, d) Proteine, welche das Template stabilisieren und gegebenenfalls e) Proteine, welche die Polymerase an das Template führen .

Die unter b) genannten Proteine bilden Strukturen aus, die entweder offen oder geschlossen sind, beispielsweise ringförmige oder halbringförmige Strukturen. Derlei Strukturen können durch eine oder mehrere Spezies von Proteinen ausgebildet werden. Dabei ist es möglich, daß eine der besagten Proteinspezies eine Polymeraseaktivität aufweist.

Die für die Ausbildung dieser Strukturen verantwortlichen Proteine werden, sofern sie keine Polymeraseaktivität aufweisen, hierin im folgenden als „Gleitklammerproteine“ oder „Klammerproteine“ bezeichnet.

5

Beispielhaft sei für einen prozessiven Replikationsapparat, der nicht geschlossen ringförmig ist, auf diejenigen des Bakteriophagen T4 oder T7 hingewiesen.

- 10 Beispielhaft sei für einen prozessiven Replikationsapparat, der geschlossen ringförmig ist, auf diejenigen des Bakteriums *E. coli* hingewiesen (Stukenberg, P.T., Studwell-Vaughan, P.S. & O'Donell, M. (1991) Mechanism of the β -clamp of DNA polymerase III holoenzyme. J. Biol. Chem. 266, 11328-11334; Kuriyan, J. & O'Donnel, M. (1993) Sliding clamps of DNA polymerases. J.Mol.Biol.234, 15 915-925).

- Es ist bekannt, daß der Replikationsapparat in Archaeobakterien dem eukaryontischen Replikationsapparat ähnlich ist, obwohl die Genomorganisation in Eukaryonten und Archaeobakterien gänzlich verschieden ist und die zelluläre Struktur der Eubakterien dem der Archaea ähnelt. (Edgell, D.R. and Doolittle, W.F. (1997). Archaea and the origin(s) of DNA replication proteins. Cell 89, 995-998.).

- Die Gleitklammer ist häufig über ein oder mehrere weitere Proteine an ein Elongationsprotein gebunden, mit anderen Worten an das Elongationsprotein gekoppelt. Ein derartiges koppelndes Protein wird hierin im folgenden als Kopplungsprotein oder koppelnde Untereinheit bezeichnet, wobei gegebenenfalls die Kopplung über eine Mehrzahl von Kopplungsproteinen erfolgen kann.

- 30 Unter Elongationsprotein soll hierin im übrigen ein Polymeraseaktivität-aufweisendes Protein oder Komplex verstanden werden, das oder der

mindestens eine oder mehrere der folgenden Eigenschaften aufweist:
Verwendung von RNA als Template zur Synthese von DNA und/oder RNA,
Verwendung von DNA als Template zur Synthese von DNA und/oder RNA,
Synthese von RNA, Synthese von DNA, Synthese von Nukleinsäuren aus DNA
5 und RNA, Exonukleaseaktivität in 5'-3'- Richtung und Exonukleaseaktivität in
3'-5'-Richtung, Strangverdrängungsaktivität (strand-displacement activity)
Thermostabilität und Prozessivität oder Nicht-Prozessivität.

Die dreidimensionale Struktur von verschiedenen Gleitklammerproteinen wurde
10 bereits bestimmt:

- die des eukaryontischen proliferating cell nuclear antigen (PCNA) (Krishna, T.S.R., Kong, X.-P., Gary, S., Burgers, P. M. & Kuriyan, J. (1994) Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. Cell 79, 1233-1243; Gulbis, J. M., Kelman, Z., Hurwitz, J., O'Donnel, M. & Kuriyan, J. (1996) Structure of the C-terminal region of p21WAF1/CIP1 complexed with human PCNA. Cell 87, 297-306),
15
- die der β -Untereinheit der Polymerase III des Eubakteriums *Escherichia coli* (Kong, X.-P., Onrust, R., O'Donnell, M. & Kuriyan, J. (1992) Three dimensional structure of the β -subunit of *Escherichia coli* DNA polymerase III
20 holoenzyme; a sliding DNA clamp. Cell 69, 425-437)
- und die des Gen45-Proteins des Bakteriophagens T4 Proteins (Kelman, Zvi, Hurwitz, J. O'Donnel, Mike (1998) Structure, 6, 121-125).

Die Gesamtstruktur dieser Gleitklammern ist sehr ähnlich; die Bilder der
25 ringförmig ausgebildeten Proteingesamtstruktur von PCNA, der β -Untereinheit und gp45-Ringe sind übereinandergelegt deckungsgleich (Kelman, Z. & O'Donnell, M. (1995) Structural and functional similarities of prokaryotic and eukaryotic sliding clamps. Nucleic Acids Res. 23, 3613-3620). Jeder Ring hat vergleichbare Dimensionen und eine zentrale Öffnung, die groß genug ist, um
30 Duplex-DNA, also einen DNA-Doppelstrang, bestehend aus den zwei komplementären DNA-Strängen, zu umschließen.

Die Gleitklammer kann sich *in vivo* nicht selbst um die DNA herum positionieren, sondern muß um die DNA fixiert werden. In Prokaryonten und Eukaryonten besteht ein derartiger Proteinkomplex aus einer Vielzahl von 5 Untereinheiten, der beim Eubakterium *Escherichia coli* γ -Komplex und beim Menschen Replikationsfaktor C (RF-C) genannt wird (Kelman, Z. & O'Donnell, M. (1994) DNA replication – enzymology and mechanisms. Curr. Opin. Gent. Dev. 4, 185-195). Der Proteinkomplex erkennt das 3'-Ende des Primers des „Primer-Template-Duplexes“ und positioniert die Gleitklammer in Anwesenheit 10 von ATP um die DNA.

Im Falle des Bacteriophagen T7 wird das gleiche Ziel, eine prozessive DNA-Synthese, mittels eines strukturell anderen Proteinkomplexes erreicht. Der Phage exprimiert eine eigene katalytische Polymerase, die T7-Polymerase, die 15 das Genprodukt des *gene 5*, welche mit einem Protein aus dem Wirt *Escherichia coli*, dem Thioredoxin, eine Bindung eingeht und als Replikase eine hochprozessive DNA-Replikation ermöglicht (Proc. Natl. Acad. Sci. U S A 1992 Oct 15; 89(20):9774-9778 Genetic analysis of the interaction between bacteriophage T7 DNA polymerase and *Escherichia coli* thioredoxin, Himawan 20 JS, Richardson CC). Auch hierbei kommt es zur Klammerbildung, jedoch weist diese Klammer nicht die gleiche Struktur auf, wie z.B. im Falle des eukaryontischen PCNA.

Oft ist es nötig, wie zum Beispiel im Falle der humanen Polymerase δ , daß 25 Kopplungsproteine eine Verbindung zwischen dem katalytisch aktiven Teil der Polymerase und dem Prozessivitätsfaktor (Gleitklammer) zu schaffen. Beim Menschen ist dies die kleine Untereinheit der δ -Polymerase (Zhang, S.-J., Zeng, X.-R., Zhang, P., Toomey, N.L., Chuang, R.Y., Chang, L.-S., and Lee, M.Y.W.T. (1994). A conserved region in the amino terminus of DNA 30 polymerase δ is involved in proliferating cell nuclear antigen binding. J. Biol.

Chem. 270, 7988-7992). Im Falle der T7 Polymerase jedoch bindet der Prozessivitätsfaktor die katalytische Einheit der Polymerase direkt.

DNA-Polymerasen werden unter anderem durch zwei Eigenschaften
5 charakterisiert, ihre Elongationsrate, das heißt die Anzahl der Nukleotide, die sie pro Sekunde in einen wachsenden DNA-Strang inkorporieren können, und ihre Dissoziationskonstante. Wenn die Polymerase nach jedem Inkorporationsschritt eines der Nukleotide in die wachsende Kette wieder vom Strang abdissoziiert (d.h. ein Elongationsschritt erfolgt pro Bindungsereignis),
10 dann hat die Prozessivität den Wert 1 und die Polymerase ist nicht prozessiv. Diese Synthese wird distributiv genannt. Wenn die Polymerase für wiederholte Nukleinsäureinkorporationen mit dem Strang verbunden bleibt, dann wird der Replikationsmodus als prozessiv bezeichnet und kann einen Wert von mehreren Tausend erreichen (siehe hierzu auch: Methods in Enzymology
15 Volume 262, DNA Replication, Edited by J. L. Campbell, Academic press 1995, pp. 270-280)

Für die meisten *in vitro* Anwendungen, wie PCR oder Sequenzierungsprozesse, ist Prozessivität eine wünschenswerte Eigenschaft,
20 die allerdings die bislang in diesen Reaktionen eingesetzten thermostabilen Enzyme nur in geringem Maße besitzen, wohingegen die temperaturempfindlichen, mit Thioredoxin assoziierte T7 Polymerase eine Prozessivität von einigen tausend Nukleotiden aufweist. Im Vergleich - die thermostabile DNA-Polymerase aus *Thermus thermophilus* oder *Thermus*
25 *aquaticus* haben nur eine Prozessivität von etwa 50 Nukleotiden (Biochim Biophys Acta 1995 Nov 7;1264(2):243-248 Inactivation of the 5'-3' exonuclease of *Thermus aquaticus* DNA polymerase. Merken LS, Bryan SK, Moses RE).

Die U.S. Patente 4,683,195, 4,800,195 und 4,683,202 beschreiben die
30 Anwendung solcher thermostabiler DNA-Polymerasen in der Polymerase-Kettenreaktion (PCR). In der PCR wird unter Verwendung von Primern,

Template (auch Matrize genannt), Nukleotiden, einer DNA-Polymerase eines entsprechenden Puffers und geeigneten Reaktionsbedingungen DNA neu synthetisiert. Hierbei wird üblicherweise von einer doppelsträngigen DNA-Sequenz ausgegangen, von der ein bestimmter Zielbereich amplifiziert werden soll. Hierbei werden zwei Primer eingesetzt, welche zu flankierenden Bereichen der Zielsequenz auf jeweils einem Teilstrang des DNA-Doppelstranges komplementär sind. Zur Anhybridisierung der Primer werden allerdings die DNA-Doppelstränge zuerst denaturiert, insbesondere thermisch aufgeschmolzen. Nach Anhybridisierung der Primer erfolgt eine Elongation mittels der Polymerase. Daraufhin wird nochmals denaturiert und damit die neu gebildeten DNA-Stränge von den Template-Strängen getrennt, woraufhin für einen weiteren Elongationszyklus neben den ursprünglichen Template-Strängen auch die im ersten Schritt gebildeten Nukleinsäure-Stränge als Template zur Verfügung stehen, diese jeweils erneut mit Primern hybridisiert werden und eine erneute Elongation stattfindet. Diese Vorgehensweise wird zyklisch durchgeführt unter jeweils thermischer Denaturierung als Zwischenschritte. Bei der PCR kommt es bevorzugt zur Verwendung einer thermostabilen Polymerase, welche das zyklische, thermische Aufschmelzen der DNA Stränge übersteht. So wird häufig *Taq* DNA-Polymerase verwendet (U.S. Patent 4,965,188). Die Prozessivität der *Taq* DNA-Polymerase ist jedoch, wie oben ausgeführt, relativ gering im Vergleich zur T7-Polymerase.

DNA-Polymerasen finden auch bei der DNA-Sequenzbestimmung Anwendung (Sanger et al., *Proc. Natl. Acad. Sci.*, USA 74:5463-5467 (1997)). Häufig wird bei der Sequenzierung nach Sanger eine T7-DNA-Polymerase verwendet (Tabor, S. und Richardson, C.C. *Proc. Natl. Acad. Sci.*, USA 86:4076-4080 (1989)). Später wurde das Cycle-Sequencing-Verfahren entwickelt (Murray, V. (1989) *Nucleic Acids Res.* 17, 8889), welches kein einzelsträngiges Template erfordert und die Initiierung der Sequenzreaktion mit verhältnismäßig geringen Mengen an Template erlaubt. Die hierbei zur Verwendung kommenden Polymerasen können z.B. die oben erwähnte *Taq*-Polymerase sein (U.S.

Patent 5,075, 216) oder die Polymerase von *Thermotoga neapolitana* (WO 96/10640) oder andere thermostabile Polymerasen. Neuere Verfahren koppeln die exponentielle Amplifikation und die Sequenzierung eines DNA-Fragmentes in einem Schritt, so daß es möglich ist, genomische DNA direkt zu 5 sequenzieren. Eines der Verfahren, das sogenannte DEXAS-Verfahren (Nucleic Acids Res 1997 May 15;25(10):2032-2034 Direct DNA sequence determination from total genomic DNA. Kilger C, Pääbo S, Biol. Chem. 1997 Feb; 378(2):99-105 Direct exponential amplification and sequencing (DEXAS) of genomic DNA. Kilger C, Pääbo S und DE 19653439.9 sowie DE 10 19653494.1), verwendet eine Polymerasen mit verminderter Diskriminierungsfähigkeit gegenüber Dideoxynukleotiden (ddNTPs) im Vergleich zu Deoxynukleotiden (dNTPs) sowie einen Reaktionspuffer, zwei Primer, die bevorzugt nicht äquimolar vorliegen, und die oben genannten Nukleotide, um dann in mehreren Zyklen eine komplette, sequenzspezifische 15 DNA-Leiter eines Fragmentes zu erhalten, welches von den Primern flankiert ist. Eine Weiterentwicklung dieses Verfahrens besteht in der Verwendung eines Polymerasegemisches, wobei eine der beiden Polymerasen zwischen ddNTPs und dNTPs diskriminiert, während die zweite eine verminderte Diskriminierungsfähigkeit aufweist (Nucleic Acids Res. 1997 May 15; 20 25(10):2032-2034 Direct DNA sequence determination from total genomic DNA. Kilger C, Pääbo S).

DNA-Polymerasen finden auch Anwendung bei der reversen Transkription von RNA in DNA. Hierbei dient RNA als Template und die Polymerase synthetisiert 25 einen komplementären DNA-Strang. Zur Anwendung kommt hier z.B. die thermostabile DNA-Polymerase aus dem Organismus *Thermus thermophilus* (Tth) (U.S. Patent 5,322,770).

Es kann auch vorkommen, daß die Polymerase eine 'proof-reading' Aktivität 30 besitzt, also eine 3'-5' Exonukleaseaktivität aufweist. Diese Eigenschaft ist insbesondere dann wünschenswert, wenn das zu synthetisierende Produkt mit

einer niedrigen Fehlerrate bei der Nukleotidinkorporation hergestellt werden soll. Ein Beispiel stellen Polymerasen aus dem Organismus *Pyrococcus wosei* dar.

- 5 Die oben genannten Enzyme, die üblicherweise in PCR-Reaktionen eingesetzt werden, gehören *in vivo* größtenteils nicht zu den eigentlichen Replikationsenzymen, sondern es sind zumeist Enzyme, von denen man annimmt, daß sie bei der DNA-Reparatur beteiligt sind, weshalb deren Prozessivität relativ gering ist.

10

Somit war es eine Aufgabe der vorliegenden Erfindung, mehrere der vorgenannten Eigenschaften von Polymerasen, insbesondere hohe Prozessivität und Thermostabilität, für die Verwendung in *in vitro* Reaktionen zu vereinen.

15

Diese Aufgabe wurde erfindungsgemäß gelöst durch die Bereitstellung eines thermostabilen *in vitro*-Komplexes, umfassend ein thermostabiles Gleitklammerprotein und ein thermostabiles Polymeraseaktivität-aufweisendes Elongationsprotein. Der erfindungsgemäße Komplex kann dabei zur Template-
20 abhängigen Elongation von Nukleinsäure(n) dienen.

Dieser Komplex kann in *in vitro*-Reaktionen, wie z.B. in PCR-Reaktionen, eingesetzt werden und weist dabei eine hohe Prozessivität auf. Von Vorteil ist zudem, daß der Komplex eine geringe Fehlerquote bei der
25 Nukleotidinkorporation aufweist, also eine erhöhte Genauigkeit beim Baseneinbau hat. Dieser Komplex kann somit bei der Elongation, der Amplifikation, der Reversen Transkription, DNA-Markierung und der Sequenzierung von Nukleinsäuren in vorteilhafter Weise eingesetzt werden. Die vorstehend aufgezeigten Verwendungsmöglichkeiten stellen jeweils für sich
30 besonders bevorzugte Ausführungsformen der Erfindung dar.

Im Falle der Verwendung des erfindungsgemäßen Komplexes zur Amplifikation von Nukleinsäuren wurde überraschend festgestellt, daß das dabei hergestellte Amplifikationsprodukt eine besonders niedrige Basenfehlinkorporationsrate aufweist.

5

Bei der Verwendung eines solchen Komplexes, beispielsweise in Standard-PCR-Reaktionen, wird eine einfache Handhabung und eine hohe Prozessivität gewährleistet, wie dies beispielsweise aus Fig. 26 ersichtlich ist.

10 In einer Ausführungsform ist vorgesehen, daß bei dem erfindungsgemäßen *in vitro*-Komplex das Gleitkammerprotein mit dem Elongationsprotein verbunden ist.

In einer bevorzugten Ausführungsform des erfindungsgemäßen thermostabilen
15 *in vitro*-Komplexes sind das Gleitkammerprotein und das Elongationsprotein über ein Kopplungsprotein verbunden.

Die Kopplung zwischen Gleitkammerprotein und Polymeraseaktivität aufweisendem Elongationsprotein kann durch kovalente, aber auch durch
20 nicht-kovalente Bindung erfolgen. In einer bevorzugten alternativen Ausführungsform ist vorgesehen, daß eine direkte Kopplung zwischen Gleitkammerprotein und Elongationsprotein erfolgt.

Dabei kann vorgesehen sein, daß das Gleitkammerprotein und/oder das
25 Elongationsprotein aus Archaeobakterien stammen.

In einer bevorzugten Ausführungsform des erfindungsgemäßen Komplexes handelt es sich um einen prokaryontischen *in vitro*-Komplex.

In einer bevorzugten Ausführungsform kann es sich bei dem erfindungsgemäßen prokaryontischen Komplex um einen archaebakteriellen *in vitro*-Komplex handeln.

5 In einer bevorzugten alternativen Ausführungsform kann es sich bei dem erfindungsgemäßen prokaryontischen Komplex um einen eubakteriellen *in vitro*-Komplex handeln.

In einer alternativen Ausführungsform des erfindungsgemäßen Komplexes
10 handelt es sich um einen eukaryontischen *in vitro*-Komplex.

In diesem Zusammenhang ist ein prokaryontischer *in vitro*-Komplex ein solcher, bei dem das Gleitklammerprotein prokaryontischen Ursprungs ist unabhängig vom Ursprung des Elongationsproteins. Entsprechend ist ein
15 eubakterieller Komplex ein solcher, bei dem das Gleitklammerprotein eubakteriellen Ursprungs ist, unabhängig vom Ursprung des Elongationsproteins. Entsprechend ist ein archaebakterieller Komplex ein solcher, bei dem das Gleitklammerprotein archaebakteriellen Ursprungs ist, unabhängig vom Ursprung des Elongationsproteins. Entsprechend ist weiterhin
20 ein eukaryontischer Komplex ein solcher, bei dem das Gleitklammerprotein eukaryontischen Ursprungs ist, unabhängig vom Ursprung des Elongationsproteins.

Gegenstand der vorliegenden Erfindung sind ebenfalls solche thermostabilen *in vitro*-Komplexe, bei denen die Komplexe aufbauenden Proteine zum Teil
25 aus Archaeobakterien, Eukaryonten und aus Eubakterien stammen. Insoweit sind jegliche Permutationen der *in vitro*-Komplexe hinsichtlich der Proteinkomponenten und ihres jeweiligen Ursprunges Gegenstand der vorliegenden Erfindung.

Ursprung im vorstehenden Sinne soll dabei jene Quelle bezeichnen, auf die das Gen, die Geninformation oder das Protein zurückzuführen ist.

Davon unabhängig ist die tatsächliche Art und Weise der Gewinnung des
5 Gleitklammerproteins bzw. des Elongationsproteins, die beispielsweise durch chemische Synthese, gentechnologische Verfahren oder Isolierung aus den natürlichen Quellen erfolgen kann.

Insbesondere ist Gegenstand der Erfindung ein thermostabiler
10 prokaryontischer *in vitro*-Komplex zur Elongation, insbesondere zur Template-abhängigen Elongation von Nukleinsäuren, der ein thermostabiles Gleitklammerprotein (Fig. 20), welches die komplementären Nukleinsäurestränge ganz oder teilweise umschließt, und ein thermostabiles, Polymeraseaktivität (Fig. 21 und Fig. 22) aufweisendes Protein umfaßt, wobei
15 dieses Protein oder dieser Proteinkomplex mit dem Gleitklammerprotein gekoppelt oder assoziiert ist.

Im Rahmen der vorliegenden Erfindung umfaßt der Begriff Polymeraseaktivität-aufweisendes Elongationsprotein auch Polymeraseaktivität-aufweisende
20 Proteinkomplexe oder Untereinheiten solcher Komplexe, welche die Polymeraseaktivität tragen.

Thermostabil im Sinne der vorliegenden Erfindung bedeutet, daß der *in vitro*-Komplex mit hoher Prozessivität Nukleotide in wachsende Nukleinsäurestränge
25 inkorporiert, sowohl bei niedrigen als auch bei hohen Temperaturen, die in der PCR oder einer anderen Reaktion auftreten, wie z.B. der DNA-Sequenzierung.

Die PCR besteht z.B. in der Regel aus den Schritten der Denaturierung (70°C bis 98°C), dem Annealing (40°C bis 78°C) und der DNA-Strangsynthese (60°C
30 bis 76°C). Somit muß dieser Komplex mindestens zwischen ca. 60°C und ca. 70°C, bevorzugt insbesondere zwischen 60°C und 76°C und besonders

bevorzugterweise funktionsfähig sein zwischen 40°C und 98°C. Es dürfen während der gesamten Reaktion keine irreversiblen Denaturierungserscheinungen des Komplexes oder einzelner Komponenten auftreten, welche die Elongationsreaktion unterbinden oder inhibieren.

5

Die Gleitklammer

Die folgenden Ausführungen sollen dazu dienen, die Funktion und die möglichen Erscheinungsformen der Gleitkammer bzw. des Gleitklammerproteins zu veranschaulichen.

10

Das Gleitklammerprotein erfüllt die Funktion, das Elongationsprotein an die DNA zu binden. Wie eingangs bereits ausgeführt, umschließt das Gleitklammerprotein selbst die einzelsträngige oder doppelsträngige Nukleinsäure ganz oder teilweise oder durch Assoziation an das Polymeraseaktivität aufweisende Protein beziehungsweise an den Polymeraseaktivität-aufweisenden Proteinkomplex oder eine Untereinheit davon und bildet dadurch eine Klammer um die Nukleinsäure aus. In jedem Fall wird durch diese Klammerbildung die Prozessivität signifikant gesteigert, mindestens um das Eineinhalbfache (Beispiel 5 und Fig. 23).

20

Das heißt, der erfindungsgemäße *in vitro*-Komplex besitzt eine mindestens eineinhalbfache Prozessivität im Vergleich zu einem Elongationsprotein alleine, beziehungsweise im Vergleich zu einem Polymeraseaktivität-aufweisenden Proteinkomplex, bzw. einer Untereinheit davon, ohne Gleitklammer (Beispiel 5 und Fig. 23).

Als Gleitkammer können erfindungsgemäß beispielsweise Homologe oder Funktionsanalogue des "Proliferating Cell Nuclear Antigen"-Proteinkomplexes, kodiert im menschlichen Genom, oder Homologe des ebenfalls ringförmigen "β-clamp"-Proteinkomplex aus *E. coli* dienen, welche aus thermostabilen Organismen stammen und thermostabil sind, oder sofern sie nicht-thermostabil

30

sind, thermostabil gemacht werden, oder aus nicht thermostabilen Organismen stammen und thermostabil sind bzw. nachträglich durch Veränderung der Aminosäuresequenz thermostabil gemacht werden (Eijsink VG, van der Zee JR, van den Burg B, Vriend G, Venema G, FEBS Lett 1991 Apr 22;282(1):13-16, Improving the thermostability of the neutral protease of *Bacillus stearothermophilus* by replacing a buried asparagine by leucine, Bertus Van den Burg, Gert Vriend, Oene R. Veltman, Gerard Venema, and Vincent G. H. Eijsink Engineering an enzyme to resist boiling PNAS 1998 95: 2056-2060). Unter homologen Sequenzen soll hierin im folgenden Sequenzen verstanden werden, die sich dadurch auszeichnen, daß sie zu einer oder mehreren anderen Sequenzen eine Sequenzähnlichkeit aufweisen und zwar in einem Maße, bei dem nicht von einer zufälligen Ähnlichkeit ausgegangen werden kann. Der Grad der Sequenzähnlichkeit wird in Prozent ausgedrückt und Homologie genannt. Bisweilen wird auch der Begriff „Sequenzidentität“ verwendet. Ein Homolog ist eine Nukleinsäure oder Aminosäuresequenz, welche zu einer Bezugssequenz eine homologe Sequenz ist.

Die Gleitklammer kann aus mehreren Komponenten aufgebaut sein. Die im menschlichen Genom identifizierte Gleitklammer besteht aus drei PCNA-Protein Komponenten (SEQ ID NO: 11) (Homotrimeres), die im *E. coli*-Genom identifizierte Gleitklammer besteht aus zwei Komponenten (SEQ ID NO: 35) (Homodimeres).

Als Gleitklammer im Sinne der vorliegenden Erfindung ist hierbei insbesondere jedes Protein zu verstehen, das die funktionelle Eigenschaft der Polymeraseprozessivitätssteigerung (Beispiel 5 und Fig. 23) besitzt und/oder der Senkung der Fehlerate dient. Dazu kann die Gleitklammer eine ringförmige dreidimensionale Struktur aufweisen oder durch Kopplung an ein anderes Protein eine ringförmige dreidimensionale Struktur bilden, durch die sie in der Lage ist, ein- und doppelsträngige Nukleinsäuren ganz oder teilweise zu umschließen.

Als Gleitklammer im Sinne der vorliegenden Erfindung ist insbesondere ein solches Protein zu verstehen, das

- 5 1. zu der menschlichen PCNA-Aminosäuresequenz (Eukaryonten) (SEQ ID NO 11) auf einer Länge von mindestens 100 Aminosäuren bei einem Sequenzalignment eine mindestens 20 %ige, bevorzugt mindestens 25 %ige und noch bevorzugter mindestens 30 %ige Sequenzidentität aufweist oder das
 - 10 2. zu der bakteriellen β -clamp-Sequenz aus *E. coli* (Eubakterien) (SEQ ID NO 35) auf einer Länge von mindestens 100 Aminosäuren bei einem Sequenzalignment eine mindestens 20 %ige, bevorzugt mindestens 25 %ige und noch bevorzugter mindestens 30 %ige Sequenzidentität aufweist oder das
 - 15 3. zu der Aminosäuresequenz des PCNA-Homologen aus *Archaeoglobus fulgidus* (Archaeobakterien) (SEQ ID NO 12) auf einer Länge von mindestens 100 Aminosäuren bei einem Sequenzalignment eine mindestens 20 %ige, bevorzugt 25 %ige und noch bevorzugter 30 %ige Sequenzidentität aufweist.
 - 20
- Sämtliche Sequenzalignments, wie sie hierin offenbart werden, wurden mit dem BLAST Algorithmus nach Altschul, S.F., Gish, W. Miller, W., Myers, E.W., and Lipman, D.J., J. Mol. Biol. 215, 403-410 (1990) erstellt.

- 25 Die erfindungsgemäße Gleitklammer kann eines oder mehrere der vorgenannten Merkmale aufweisen.

Als Gleitklammer bzw. Gleitklammerproteine im Sinne der vorliegenden Erfindung sind außerdem Proteine zu verstehen, die eine oder beide der
30 folgenden Konsensussequenzen (Region 1 und Region 2) beinhalten und an

nicht mehr als vier Positionen von der Region 1 (SEQ ID No.: 39) oder nicht mehr als vier Positionen von der Region 2 (SEQ ID No.: 40) abweichen (Fig. 4):

Region 1

5 (SEQ ID No.: 39):

[GAVLIMPFW]-D-X-X-X-[GAVLIMPFW]-X-X-[GAVLIMPFW]-X-[GAVLIMPFW]-
X-[GAVLIMPFW]-X-X-X-X-F-X-X-Y-X-X-D

und / oder

Region 2:

10 (SEQ ID No.: 40):

[GAVLIMPFW]-X(3)-L-A-P-[KRHDE]-[GAVLIMPFW]-E

Die Aminosäuren werden hierbei gemäß der Standard IUPAC - Einbuchstaben - Nomenklatur benannt und gemäß dem Prosite Pattern-Beschreibungsstandard aufgeführt. Dabei sind die folgenden Aminosäuregruppen häufig zusammengefaßt:

G,A,V,L,I,M,P,F oder W (Aminosäuren mit nicht polaren Seitenketten)

S,T,N,Q,Y, oder C (Aminosäure mit ungeladenen polaren Seitenketten)

K,R,H,D oder E (Aminosäure mit geladenen und polaren Seitenketten)

20 Außerdem bedeutet X in den Sequenzen bzw. Sequenzprotokollen jede beliebige Aminosäure oder Insertion oder Deletion

Aus dem in Fig. 12 dargestellten multiplen Alignment von menschlichen PCNA-Homologen wurde zudem ein Hidden-Markov-Modell generiert. Als
25 Gleitklammer im Sinne der vorliegenden Erfindung ist somit insbesondere auch jedes Protein zu verstehen, das mit dem so erzeugten Hidden-Markov-Modell (hierin im folgenden mit „HMM“ bezeichnet) einen „Score“ von mehr als 20, bevorzugt 25 am bevorzugtesten 30 aufweist (Fig. 12), wobei ein „Score“ der Ausgabewert einer HMM-Analyse ist. Die Hidden-Markov-Modelle und die
30 entsprechenden Scores wurden mit dem hmmfs Programm (Version 1.8.4, July 1997) aus dem HMMER-Paket berechnet (HMMER Protein and DNA Hidden

Markov Models (Version 1.8) von Sean Eddy, Dept. of Genetics, Washington University School of Medicine, St. Louis, USA;).

Markov-Modelle mit verstecktem Profil (engl.: profile hidden Markov models –
5 profile HMMs), kurz auch „Hidden Markov Models“ genannt, hierin als HMM
abgekürzt, sind statistische Modelle des Konsensus der Primärstruktur einer
Sequenzfamilie. Die Profile verwenden positionsspezifische Punktzahlen
(„scores“) für Aminosäuren (oder Nukleotide) und positionsspezifische
Punktzahlen zum Eröffnen oder Erweitern einer Insertion oder Deletion.
10 Methoden zum Erstellen von Profilen, ausgehend von multiplen Alignments,
wurden von Taylor (1986), Gribskov et al. (1987), Barton (1990) und
Heinikoff (1996) eingeführt.

HMM stellen eine vollkommen probabilistische Beschreibung von Profilen
15 bereit, d. h. die Lehre von Bayes legt fest, wie die gesamten
Wahrscheinlichkeits- (Auswertungs-) Parameter gesetzt werden sollten
(vergl. Krogh et al. 1994, Eddy 1996 und Eddy 1998). Die zentrale Idee ist
die, daß ein HMM ein finites Modell ist, das eine
Wahrscheinlichkeitsverteilung über eine unbegrenzte Anzahl möglicher
20 Sequenzen beschreibt. Das HMM setzt sich aus einer Anzahl von Zuständen
zusammen, die den Säulen eines multiplen Alignments, wie dies
üblicherweise dargestellt wird, entsprechen. Jeder Zustand emittiert Symbole
(Reste) entsprechend den (zustandsspezifischen) Symbol-Emissions-
Wahrscheinlichkeiten, und die Zustände sind durch Zustand-Übergangs-
25 Wahrscheinlichkeiten untereinander verbunden. Ausgehend von einem
Anfangszustand wird eine Abfolge von Zuständen generiert, indem von
einem Zustand zum anderen entsprechend den Zustands-Übergangs-
Wahrscheinlichkeiten übergegangen wird, bis ein Endzustand erreicht ist.
Jeder Zustand emittiert dann Symbole entsprechend den Emissions-
30 Wahrscheinlichkeits-Verteilung dieses Zustands, was eine beobachtbare
Sequenz von Symbolen erzeugt.

Das Attribut „hidden“ (versteckt) leitet sich ab von der Tatsache, daß die zugrunde liegende Zustandssequenz nicht beobachtet werden kann; was gesehen wird, ist die Symbolsequenz. Eine Abschätzung der Übergangs- und Emissions-Wahrscheinlichkeiten (das Trainieren des Modells) wird durch dynamische Programmierungsalgorithmen erreicht, die in dem HMMER-Paket implementiert sind.

Wenn ein existierendes HMM und eine Sequenz gegeben sind, kann man die Wahrscheinlichkeit berechnen, daß das HMM die fragliche Sequenz erzeugen könnte. Das HMMER-Paket liefert eine numerische Größe (die Punktzahl bzw. Ausgabewert), die proportional zu dieser Wahrscheinlichkeit ist, d. h. den Informationsgehalt der Sequenz, angegeben in Bits, gemessen gemäß dem HMM.

15

In Zusammenhang mit HMM sei auf folgende Literaturstellen verwiesen:

Barton, G.J. (1990):

Protein multiple alignment and flexible pattern matching.

20 Methods Enzymol. 183: 403-427.

Eddy, S.R. (1996):

Hidden markov models.

Curr. Opin. Struct. Biol. 6: 361-365.

25

Eddy, S.R. (1998):

Profile hidden markov models.

Bioinformatics. 14: 755-763.

30

Gribskov, M. McLachlan, A.D. und Eisenberg D. (1987):

Profile analysis: Detection of distantly related proteins.

Proc. Natl. Acad. Sci. USA. 84: 4355-5358.

Heinikoff, S. (1996):

Scores for sequence searches and alignment.

5 Curr. Opin. Strct. Biol. 6: 353-360.

Krogh, A., Brown, M., Mian, I.S., Sjolander, K. und Haussler, D.
(1994):

10 Hidden markov models in computational biology: Applications to
protein modelling.

J. Mol. Biol. 235: 1501-1531.

Taylor, W.R. (1986):

15 Identification of protein sequence homology by consensus template
alignment.

J. Mol. Biol. 188: 233-258.

Aus dem in Fig. 13 dargestellten multiplen Alignment von *E. coli* β -clamp-Homologen wurde ein HMM generiert. Als Gleitklammer im Sinne der
20 vorliegenden Erfindung ist somit insbesondere auch jedes Protein zu verstehen, das mit dem so erzeugten HMM einen Score von mehr als 25, bevorzugt 30 und am bevorzugtesten 35 aufweist (Fig. 13).

Die Gleitklammer kann aus mehreren Komponenten aufgebaut sein, die durch
25 eine charakteristische Bindung fest aneinander gebunden sind, so daß ein stabiler ringförmiger Molekülkomplex gebildet wird, der nicht ohne weiteres von der Nukleinsäure dissoziieren kann. Dadurch wird eine feste, aber nicht kovalente Bindung an die Nukleinsäure ermöglicht, die die freie Verschiebbarkeit auf derselben aber nicht behindert. Die
30 prozessivitätssteigernden Gleitklammerproteine haben zudem charakteristische lokale Moleküleigenschaften im Bereich der Wechselwirkungsregion zur DNA,

die die freie Verschiebbarkeit erleichtern und die durch in diese Region eingelagerte Wassermoleküle unterstützt werden kann.

Eine bevorzugte Ausführungsform der vorliegenden Erfindung ist weiterhin
5 insbesondere ein thermostabiler prokaryontischer *in vitro*-Komplex, wobei das Gleitklammerprotein eines der folgenden ist: AF 0335 aus *Archaeoglobus fulgidus* (SEQ ID NO.: 12) (Fig.24), MJ0247 aus *Methanococcus jannaschii* (SEQ ID NO.: 13), PHLA008 aus *Pyrococcus horikoshii* (SEQ ID NO.: 14), MTH1312 aus *Methanobacterium Thermoautotrophicus* (SEQ ID NO.: 15),
10 sowie AE000761_7 aus *Aquifex aeolicus* (SEQ ID NO.: 36).

Insbesondere sind solche thermostabilen *in vitro*-Komplexe Gegenstand dieser Anmeldung, wobei das Gleitklammerprotein eine Aminosäuresequenz umfaßt, die ausgewählt ist aus der Gruppe, die SEQ ID No. 11, 12, 13, 14, 15 und 36
15 umfaßt.

Der Gleitklammerlader

Desweiteren ist als bevorzugte Ausführungsform zu verstehen, wenn der erfindungsgemäße Komplex einen Gleitklammerlader umfaßt. Als
20 Gleitklammerlader wird hierin ein Protein, Proteinkomplex oder Untereinheit eines Proteins verstanden, welches ein Homolog des im Menschen identifizierten "Replication Factor C"-Proteinkomplexes umfaßt.

Dieser Proteinkomplex besteht im Menschen aus fünf Untereinheiten wobei sie
25 durch 5 separate Gene im Menschen kodiert sind. Die 4 von jeweils einem Gen kodierten kleinen Untereinheiten (hierin als Gleitklammerlader 1 bezeichnet) bilden beim Menschen einen Proteinkomplex. Das Protein der großen Untereinheit wird beim Menschen durch ein Gen kodiert (hierin als Gleitklammerlader 2 bezeichnet). Die Sequenzen der vier kleinen
30 Untereinheiten sind als SEQ ID NO.: 1, 32, 33, 34 dargestellt, die Sequenz der großen Untereinheit ist als SEQ ID NO.: 6 dargestellt.

Als Gleitklammerlader 1 können erfindungsgemäß Homologe oder Funktionsanaloge, einzeln oder in beliebiger Kombination, einer jeden der vorstehend genannten Sequenzen SEQ ID NO.: 1, 32, 33, 34 verwendet 5 werden. Dabei können die Homologen prokaryontischen, wie eubakteriellen oder archaebakteriellen, oder eukaryontischen Ursprungs sein.

Als Gleitklammerlader 2 können erfindungsgemäß Homologe oder Funktionsanaloge, einzeln oder in beliebiger Kombination, der vorstehend 10 genannten Sequenz SEQ ID NO.: 6 verwendet werden. Dabei können die Homologen prokaryontischen, wie eubakteriellen oder archaebakteriellen, oder eukaryontischen Ursprungs sein.

Als Gleitklammerlader 1 im Sinne der vorliegenden Erfindung kann auch ein 15 solches Protein verstanden werden, das zu der menschlichen (Eukaryonten) Aminosäuresequenz (SEQ ID NO: 1, 32, 33, 34) auf einer Länge von mindestens 100 Aminosäuren bei einem Sequenzalignment eine mindestens 20 %ige, bevorzugt mindestens 25 %ige und noch bevorzugter mindestens 30 %ige Sequenzidentität aufweist.

20

Als Gleitklammerlader 2 im Sinne der vorliegenden Erfindung kann auch ein solches Protein verstanden werden, das zu der menschlichen (Eukaryonten) Aminosäuresequenz (SEQ ID NO: 6) auf einer Länge von mindestens 150 Aminosäuren bei einem Sequenzalignment eine mindestens 20 %ige 25 Sequenzidentität, bevorzugt mindestens 25 %ige und noch bevorzugter mindestens 30 %ige aufweist.

Gleitklammerlader-Homologe im Sinne der obigen Definition sind beispielsweise die in Fig. 1 aufgeführten Gene aus Archaeobakterien. Diesen 30 Genen entsprechen für den Gleitklammerlader 1 die Sequenzen SEQ ID NO:

2, 3, 4, und 5 und für den Gleitklammerlader 2 die Sequenzen SEQ ID NO: 7, 8, 9, und 10.

Als Gleitklammerlader 1 im Sinne der vorliegenden Erfindung kann auch ein Protein verstanden werden, das eine Sequenz gemäß der folgenden Konsensussequenz umfaßt und an nicht mehr als vier Positionen von dieser Sequenz abweicht (Alinierung siehe auch Fig. 6):

SEQ ID No. 41:

10 C-N-Y-X-S-[KRHDE]-I-I-X-[GAVLIMPFW]-[GAVLIMPFW]-Q-S-R-C-X-X-F-R-F-X-P-[GAVLIMPFW]

Als Gleitklammerlader 2 im Sinne der vorliegenden Erfindung kann auch ein Protein verstanden werden, das eine Sequenz gemäß der folgenden Konsensussequenz umfaßt und an nicht mehr als vier Positionen von dieser Sequenz abweicht (Alinierung siehe auch Fig. 7):

SEQ ID No.42:

K-X-X-L-L-X-G-P-P-G-X-G-K-T-[STNQYC]-X-[GAVLIMPFW]-X-X-
20 [GAVLIMPFW]

Aus dem in Fig. 14 dargestellten multiplen Alignment von Sequenzen des Gleitklammerladers 1 umfassend die menschliche Sequenz sowie einige dazu homologe Sequenzen aus Archaeobakterien, wurde zudem ein HMM generiert.

25 Als Gleitklammerlader 1 im Sinne der vorliegenden Erfindung ist demzufolge auch ein Protein zu verstehen, das mit dem so erzeugten HMM einen Score von mehr als 25, bevorzugt mehr als 30 und am bevorzugtesten mehr als 35 aufweist (Alinierung siehe auch Fig. 14).

30 Aus dem in Fig. 15 dargestellten multiplen Alignment von Sequenzen des Gleitklammerladers 2 umfassend die menschliche Sequenz sowie einigen dazu

homologen Sequenzen aus Archaeobakterien wurde zudem ein HMM generiert. Als Gleitklammerlader 2 im Sinne der vorliegenden Erfindung ist demzufolge auch ein Protein zu verstehen, das mit dem so erzeugten HMM einen Score von mehr als 15, bevorzugt mehr als 20 und am bevorzugtesten mehr als 25 aufweist (Alinierung siehe auch Fig. 15).

Es kann auch vorgesehen sein, daß der erfindungsgemäße *in vitro*-Komplex ein dem Eubakterium *Escherichia coli* γ -Komplex oder Teilen davon homologes Protein als Gleitklammerlader 1 oder Gleitklammerlader 2 umfaßt.

10

Ein Gleitklammerlader im Sinne der vorliegenden Erfindung kann demgemäß ein Gleitklammerlader 1 alleine, ein Gleitklammerlader 2 alleine, oder eine Kombination einer oder mehrerer Gleitklammerlader 1 bzw. Gleitklammerlader 2, jeweils wie vorstehend definiert, sein.

15

Außerdem ist in einer Ausführungsform des erfindungsgemäßen thermostabilen *in vitro*-Komplex zur Elongation von Nukleinsäuren vorgesehen, daß dieser zusätzlich eine Verbindung umfaßt, welche bei der Spaltung Energie freisetzt, wie beispielsweise ATP, GTP, CTP, TTP, dATP, dGTP, dCTP
20 oder dTTP.

Ohne im folgenden darauf festgelegt sein zu wollen, scheint ein Gleitklammerlader die Komponenten der Gleitklammer um den ununterbrochenen DNA-Strang herum zusammenzufügen, bzw. diese wieder
25 zu entfernen, wenn die Reaktion beendet ist. Dabei kann die Gleitklammer eine ringförmige dreidimensionale Struktur aufweisen oder durch Kopplung an ein anderes Protein eine ringförmige dreidimensionale Struktur bilden, durch die sie in der Lage ist, ein- oder doppelsträngige Nukleinsäuren ganz oder teilweise zu umschließen. Der Gleitklammerlader kann dann Vorteile mit sich
30 bringen, wenn das Template-Nukleinsäuremolekül ringförmig geschlossen vorliegt.

Insbesondere sind solche thermostabile *in vitro*-Komplexe Gegenstand dieser Anmeldung, wobei der Gleitklammerlader 1 eine Aminosäuresequenz umfaßt, die ausgewählt ist aus der Gruppe, die SEQ ID No. 2, 3, 4 und 5 umfaßt.

5

Insbesondere sind solche thermostabile *in vitro*-Komplexe Gegenstand dieser Anmeldung, wobei der Gleitklammerlader 2 eine Aminosäuresequenz umfaßt, die ausgewählt ist aus der Gruppe, die SEQ ID No. 7, 8, 9 und 10 umfaßt.

10 Das Kopplungsprotein

Das Kopplungsprotein hat die Funktion, ein Elongationsprotein mit einem oder mehreren Gleitklammerproteinen oder einem Gleitklammerproteinkomplex zu verbinden. Als Kopplungsprotein im Sinne der vorliegenden Erfindung ist somit insbesondere jedes Protein zu verstehen, welches die oben beschriebene
15 Funktion aufweist. Bevorzugt sind solche Kopplungsproteine, die archaebakteriellen Ursprungs sind.

Kopplungsproteine im Sinne der vorliegenden Erfindung sind beispielsweise Homologe oder Funktionsanaloge, einzeln oder in beliebiger Kombination,
20 einer jeden der in Fig. 1 aufgeführten Sequenzen, die Homologe zu der humanen Sequenz der kleinen Untereinheit der Polymerase δ , hierin als koppelnde Untereinheit bezeichnet (Sequenzname: DPD2_HUMAN, dargestellt im Sequenzprotokoll als SEQ ID NO:16), sind. Im Sinne der vorliegenden Erfindung kann ein Kopplungsprotein ein Protein sein, das eine Sequenz
25 umfaßt, die aus der Gruppe ausgewählt ist, die die Sequenzen SEQ ID NO: 17, 18, 19, 20 und 21 umfaßt.

Als Kopplungsprotein im Sinne der vorliegenden Erfindung ist insbesondere ein solches Protein zu verstehen, das zu der menschlichen (Eukaryonten)
30 Aminosäuresequenz (SEQ ID NO: 16) auf einer Länge von mindestens 150 Aminosäuren bei einem Sequenzalignment eine mindestens 18 %ige,

bevorzugt mindestens 22 %ige und noch bevorzugter mindestens 26 %ige Sequenzidentität aufweist.

Als Kopplungsprotein im Sinne der vorliegenden Erfindung ist insbesondere ein solches Protein zu verstehen, das zu der Aminosäuresequenz (SEQ ID NO: 19) von *Pyrococcus horikoshii* auf einer Länge von mindestens 150 Aminosäuren bei einem Sequenzalignment eine mindestens 20 %ige Sequenzidentität aufweist, bevorzugt eine mindestens 25 %ige Sequenzidentität und am bevorzugtesten eine Sequenzidentität von mehr als 30 %.

10

Als Kopplungsprotein im Sinne der vorliegenden Erfindung ist auch jedes Protein zu verstehen, das die folgende Konsensussequenz umfaßt und an nicht mehr als vier Positionen von dieser Sequenz abweicht. Die Generierung der Konsensussequenz wird in Fig. 5 gezeigt, die als SEQ ID No. 43 hierin offenbart wird.

SEQ ID No. 43:

[FL]-[GAVLIMPFW]-X-X-[GAVLIMPFW]-X-G-X(13)-[GAVLIMPFW]-X-[YR]-
[GAVLIMPFW]-X-[GAVLIMPFW]-A-G-[DN]-[GAVLIMPFW]-[GAVLIMPFW]-[DS]

20

Aus dem in Fig. 16 dargestellten multiplen Alignment von Homologen zur menschlichen koppelnden Untereinheit oder Kopplungsprotein wurde zudem ein HMM generiert. Als koppelnde Untereinheit im Sinne der vorliegenden Erfindung ist somit jedes Protein zu verstehen, das mit dem so erzeugten HMM einen Score mehr als 10, bevorzugt mehr als 15, am bevorzugtesten mehr als 20 aufweist.

Insbesondere sind solche thermostabilen *in vitro*-Komplexe Gegenstand dieser Anmeldung, wobei das Kopplungsprotein eine Aminosäuresequenz umfaßt, die ausgewählt ist aus der Gruppe, die SEQ ID No. 17, 18, 19, 20 und 21 umfaßt.

30

Elongationsprotein

Im Rahmen der vorliegenden Erfindung kann ein Elongationsprotein verwendet werden, das die eingangs definierten Merkmale aufweist. Darüber hinaus können Formen von Elongationsproteinen verwendet werden, wie sie im
5 folgenden beschrieben werden und im Stand der Technik zumindest zum Teil bereits bekannt sind.

Es ist bekannt, daß einige Elongationsproteine die Anwesenheit eines Kopplungsproteins benötigen, um überhaupt Polymeraseaktivität aufzuweisen.
10 Weiterhin ist bekannt, daß einige Elongationsproteine direkt an Gleitklammerproteine binden können, wohingegen andere Elongationsproteine die Anwesenheit eines Kopplungsproteins für die Bindung an die Gleitklammerproteine benötigen. Weiterhin können Elongationsproteine die beiden vorstehend genannten Eigenschaft in sich vereinen, das heißt sowohl
15 über ein Kopplungsprotein als auch direkt, das heißt ohne ein Kopplungsprotein, an ein Gleitklammerprotein binden.

Bevorzugte Elongationsproteine für den erfindungsgemäßen *in vitro*-Komplex können ausgewählt sein aus der Gruppe, die die Organismen
20 *Carboxydotherrnus hydrogenoformans*, *Therrnus aquaticus*, *Therrnus caldophilus*, *Therrnus chliarophilus*, *Therrnus filiformis*, *Therrnus flavus*, *Therrnus oshimai*, *Therrnus ruber*, *Therrnus scotoductus*, *Therrnus silvanus*, *Therrnus species ZO5*, *Therrnus species sp. 17*, *Therrnus therrnusphilus*, *Therotoga maritima*, *Therotoga neapolitana*, *Therrnosipho africanus*,
25 *Anaerocellum therrmophilum*, *Bacillus caldotenax*, oder *Bacillus stearothermophilus* umfaßt, eingebracht wird.

Als Elongationsproteine sind beispielsweise auch die in Fig. 1 aufgeführten zum menschlichen Elongationsprotein (SEQ ID NO: 22) homologen oder
30 funktionsanalogen Elongationsproteine aus Archaeobakterien geeignet (SEQ ID NO: 23, 24, 25, 26).

Insbesondere sind solche thermostabile *in vitro*-Komplexe Gegenstand der Erfindung, wobei das Elongationsprotein eine Aminosäuresequenz umfaßt, die ausgewählt ist aus der Gruppe, die SEQ ID No. 23, 24, 25, 26, 27, 28, 29, 30 5 und 31 umfaßt.

Als Elongationsprotein im Sinne der vorliegenden Erfindung ist insbesondere ein solches Protein zu verstehen, das zu der menschlichen (Eukaryonten) Aminosäuresequenz (SEQ ID NO: 22) auf einer Länge von mindestens 200 10 Aminosäuren bei einem Sequenzalignment eine mindestens 20 %ige Sequenzidentität, bevorzugt 25 %ige Sequenzidentität, am bevorzugtesten 30 %ige Sequenzidentität aufweist.

Als Elongationsprotein im Sinne der vorliegenden Erfindung ist auch 15 insbesondere ein Protein zu verstehen, das die folgende Konsensussequenz (SEQ ID No. 44) beinhaltet und an nicht mehr als vier Positionen von dieser Sequenz abweicht. Die Fig. 8 zeigt das Alignment, das der Konsensussequenz zugrunde liegt.

20 SEQ ID No.44:

D-[GAVLIMPFW]-[GAVLIMPFW]-X-X-Y-N-X-X-X-F-D-X-P-Y-[GAVLIMPFW]-X-X-R-A

Aus dem in Fig. 17 dargestellten multiplen Alignment von Homologen zum 25 menschlichen Elongationsprotein (SEQ ID NO: 22) wurde zudem ein HMM generiert. Als Elongationsprotein im Sinne der vorliegenden Erfindung ist somit auch ein Protein zu verstehen, das mit dem so erzeugten HMM einen Score von mehr als 20, bevorzugt mehr als 25, am bevorzugtesten mehr als 30 aufweist.

Als Elongationsprotein im Sinne der vorliegenden Erfindung ist somit auch ein solches Protein zu verstehen, das zu der archaebakteriellen Aminosäuresequenz (SEQ ID NO: 27) auf einer Länge von mindestens 400 Aminosäuren bei einem Sequenzalignment eine mindestens 25 %ige Sequenzidentität, bevorzugt mindestens 30 %ige Sequenzidentität, am bevorzugtesten mindestens 35 %ige Sequenzidentität aufweist. Beispielsweise sind geeignet die aus Archaeobakterien stammenden Proteine mit der SEQ ID NO: 27, 28, 29, 30 oder 31.

10 Als Elongationsprotein im Sinne der vorliegenden Erfindung ist daher insbesondere auch jedes Protein zu verstehen, das die folgende Konsensussequenz, hierin als SEQ ID No.: 45 bezeichnet, beinhaltet und an nicht mehr als vier Positionen von dieser Sequenz abweicht. (Die Fig. 9 zeigt das Alignment, das der Konsensussequenz zu Grunde liegt.)

15

SEQ ID No. 45:

A-[GAVLIMPFW]-R-T-A-[GAVLIMPFW]-A-[GAVLIMPFW]-[GAVLIMPFW]-T-E-
G-[GAVLIMPFW]-V-X-A-P-[GAVLIMPFW]-E-G-I-A-X-V-[KRHDE]-I

20 Aus dem in Fig. 18 dargestellten multiplen Alignment von Homologen zum archaebakteriellen Elongationsprotein (SEQ ID NO: 27) wurde zudem ein HMM generiert. Als Elongationsprotein im Sinne der vorliegenden Erfindung ist somit insbesondere jedes Protein zu verstehen, daß mit dem so erzeugten HMM einen Score mehr als 35, bevorzugt mehr als 40 und noch bevorzugter mehr
25 als 45 aufweist (Fig. 18).

Das Elongationsprotein kann auch eubakteriellen Ursprungs sein.

Als Elongationsprotein im Sinne der vorliegenden Erfindung ist somit ebenfalls ein solches Protein zu verstehen, das zu der eubakteriellen
30 Aminosäuresequenz (SEQ ID NO: 37) auf einer Länge von mindestens 300 Aminosäuren bei einem Sequenzalignment eine mindestens 25 %ige,

bevorzugt mindestens 30 %ige und noch bevorzugter mindestens 35 %ige Sequenzidentität aufweist.

Als Elongationsprotein im Sinne der vorliegenden Erfindung ist daher auch insbesondere jedes Protein zu verstehen, das die folgende Konsensussequenz beinhaltet und an nicht mehr als acht Positionen von dieser Sequenz abweicht (Fig. 10):

SEQ ID No. 46:

10 [GAVLIMPFW]-P-V-G-[GAVLIMPFW]-G-R-G-S-X-[GAVLIMPFW]-G-S-
[GAVLIMPFW]-V-A-X-A-[GAVLIMPFW]-X-I-T-D-[GAVLIMPFW]-D-P-
[GAVLIMPFW]-X-X-X-[GAVLIMPFW]-L-F-E-R-F-L-N-P-E-R-[GAVLIMPFW]-S-
M-P-D

15 Aus dem in Fig. 19 dargestellten multiplen Alignment von Homologen zum eubakteriellen Elongationsprotein (SEQ ID NO: 37) wurde zudem ein HMM generiert. Als Elongationsprotein im Sinne der vorliegenden Erfindung ist somit jedes Protein zu verstehen, das mit dem so erzeugten HMM einen Score von mehr als 20, bevorzugt mehr als 25, am bevorzugtesten mehr als 30 aufweist.

20

Insbesondere sind solche thermostabile *in vitro*-Komplexe Gegenstand dieser Anmeldung, wobei das Elongationsprotein eine Aminosäuresequenz umfaßt, die ausgewählt ist aus der Gruppe, die SEQ ID No. 38 umfaßt.

25 Verwendungen des erfindungsgemäßen *in vitro*-Komplexes:

Bisher wurden DNA-Polymerasen als Elongationsproteine ohne Kopplungsprotein und ohne Gleitklammer für Standard-PCR-Reaktionen eingesetzt, so z.B. DNA-Polymerase I aus *Pyrococcus furiosus* (US No.
30 5,545,552) oder *Pyrococcus species* (EP-A-0 547 359). Diese Enzyme zeichnen sich durch die Eigenschaft aus, thermostabil zu sein und häufig eine

3'-5'-Exonuklease-Aktivität ('proof-reading' Aktivität) zu besitzen. Erst vor kurzem wurde ein Heterodimer mit Polymeraseaktivität in *Pyrococcus furiosus* entdeckt (Uemori, T., Sato, Y., Kato, I., Doi, H., and Ishino, Y. (1997). A novel DNA polymerase in the hyperthermophilic archaeon, *Pyrococcus furiosus*: 5 gene cloning, expression, and characterization. Genes to Cells 2, 499-512.)

Es ist auch möglich, durch Deletionen oder Mutationen oder durch das Anfügen von Aminosäuren die Eigenschaften dieser Proteine zu optimieren. Diese veränderten Proteine sind ebenfalls Gegenstand der vorliegenden Erfindung, 10 solange sie den erfindungsgemäßen *in vitro* Komplex zur Elongation von Nukleinsäuren bilden und die oben näher spezifizierten Funktionen erfüllen.

Fig. 3 veranschaulicht beispielhaft eine Ausführungsform des erfindungsgemäßen thermostabilen *in vitro*-Komplexes, wobei hier die 15 Gleitklammer über ein Kopplungsprotein an das Elongationsprotein bindet.

Der erfindungsgemäße *in vitro*-Komplex bzw. der diesen enthaltende erfindungsgemäße Reaktionsansatz, kann desweiteren eine Nukleinsäure umfassen, wobei die Nukleinsäure die beispielsweise zu elongierende, zu 20 sequenzierende, zu amplifizierende oder die revers zu transkribierende Nukleinsäure ist und/oder ein Primer, wobei dieser Primer bevorzugterweise an eine Nukleinsäure hybridisiert ist.

Primer sind in der Regel Oligonukleotide, welche komplementär zu einer 25 Zielsequenz sind, um an diese binden zu können, wobei sie in gegensätzlicher Orientierung, ihre 3'-Enden zueinander gerichtet, den zu elongierenden, zu sequenzierenden, zu amplifizierenden oder revers zu transkribierenden Nukleinsäureabschnitt einschließt. Sie dienen als Startpunkt der Enzymaktivität und stellen in der Regel ein freies 3'-OH Ende für die Polymerase zur 30 Inkorporation eines Nukleotides zur Verfügung.

Während der Verwendung des erfindungsgemäßen thermostabilen *in vitro*-Komplexes zur Amplifikation, Elongation, Reversen Transkription und/oder Sequenzierung liegt der erfindungsgemäße Komplex, gegebenenfalls in einem Reaktionsansatz, bevorzugterweise in einem geeigneten Puffer vor. Geeignete
5 Puffer sind solche, die für PCR, Sequenzierung, Nukleinsäuremarkierung und andere *in vitro*-Nukleinsäure-Elongationsreaktionen mittels Polymerase Anwendung finden. Geeignete Puffer werden beispielweise beschrieben in Methods in Molecular Biology, Vol. 15 Humana Press Totowa, New Jersey, 1993, Hrsg. Bruce A. White.

10

Bei der Verwendung des erfindungsgemäßen thermostabilen *in vitro*-Komplexes zur Elongation, Amplifikation, Reversen Transkription und/oder Sequenzierung kann zusätzlich zu dem erfindungsgemäßen Komplex ein Nukleotid oder ein Gemisch aus Nukleotiden vorliegen bzw. verwendet werden
15 bzw. umfaßt sein. Deoxynukleotide können aus dGTP, dATP, dTTP und dCTP ausgewählt werden, sind jedoch nicht auf diese beschränkt. Zusätzlich können auch Derivate von Deoxynukleotiden verwendet werden, die als solche Deoxynukleotide definiert sind, die in der Lage sind, durch eine thermostabile Polymerase in wachsende Nukleinsäure-Moleküle inkorporiert zu werden.
20 Solche Derivate umfassen die Thionukleotide, 7-deaza-2' -dGTP, 7-deaza-2' -dATP, Digoxigenin-dUTP (Boehringer Mannheim)-dATP sowie Deoxyinosine-Triphosphat, das auch als Ersatz-Deoxynukleotid für dATP, dGTP, dTTP oder dCTP verwendet werden kann, sind aber nicht auf diese beschränkt. Ebenso können markierte Deoxynukleotide verwendet werden. Auch die Verwendung
25 von Pyrenen und Pyrenderivaten ist möglich. Alle bekannten und/oder zu dem erfindungsgemäßen Zweck geeigneten Markierungen können dabei vorliegen bzw. verwendet werden.

Dideoxynukleotide können aus ddGTP, ddATP, ddTTP und ddCTP ausgewählt
30 werden, sind jedoch nicht auf diese beschränkt. Alternativ können auch Derivate von Dideoxynukleotiden verwendet werden, die als solche

Dideoxynukleotide definiert sind, die in der Lage sind, durch eine Polymerase in wachsende Nukleinsäure-Moleküle inkorporiert zu werden, die in der Reaktion synthetisiert werden. Solche Derivate können radioaktive Dideoxynukleotide (ddATP, ddGTP, ddTTP und ddCTP) sein oder Dideoxynukleotide (ddATP, 5 ddGTP, ddTTP und ddCTP) sein, die mit z.B. FITC, Cy5, Cy5.5, Cy7, Texas-Rot oder anderen Farbstoffen markiert sind, sind aber nicht auf diese beschränkt. Im Rahmen einer Sequenzierung unter Verwendung des erfindungsgemäßen thermostabilen *in vitro*-Komplexes können auch markierte Deoxynukleotide zusammen mit unmarkierten Dideoxynukleotiden eingesetzt 10 werden.

Ribonukleotide können aus GTP, ATP, TTP und CTP ausgewählt werden, sind jedoch nicht auf diese beschränkt. Zusätzlich können gemäß der Erfindung auch Derivate von Ribonukleotiden verwendet werden, die als solche 15 Ribonukleotide definiert sind, die in der Lage sind, durch eine Polymerase in wachsende Nukleinsäure-Moleküle inkorporiert zu werden, die in der Reaktion synthetisiert werden. Solche Derivate können radioaktive Ribonukleotide (ATP, GTP, TTP und CTP) oder Ribonukleotide (ATP, GTP, TTP und CTP), welche mit z.B. FITC, Cy5, Cy5.5, Cy7 und Texas-Rot oder anderen markiert sind, 20 umfassen, sind aber nicht auf diese beschränkt.

Während der Verwendung des Proteinkomplexes zur Amplifikation, Elongation und Sequenzierung kann es sich als vorteilhaft erweisen, wenn bei der Reaktion oder im Reaktionsansatz eine Pyrophosphatase anwesend ist. 25

Reaktionsansatz umfassend den erfindungsgemäßen thermostabilen *in vitro*-Komplex:

Weiterhin ist Gegenstand der vorliegenden Erfindung ein Reaktionsansatz, der 30 den erfindungsgemäßen *in vitro*-Komplex umfaßt. Dabei kann vorgesehen sein,

daß der Reaktionsansatz weiterhin ein oder mehrer Elongationsproteine umfaßt, die mindestens eine oder mehrere der vorstehenden Eigenschaften Aktivitäten aufweisen. Ein derartiges zusätzliches Elongationsprotein ist dabei vorteilhafterweise eine thermostabile Polymerase. Ein derartiger 5 Reaktionsansatz erlaubt eine verglichen mit der Verwendung der bekannten thermostabilen Polymerasen erhöhte Prozessivität.

**Identifizierung der Gene, Klonierung der Gene, Expression dieser und
Reinigung der Proteine der erfindungsgemäßen *in vitro*-Komplexe:**

10

In der Regel können die Komplexe, bevorzugt vollständig oder teilweise bestehend aus rekombinanten Proteinen, mit den folgenden Schritten bereitgestellt werden: Bereitstellung des Nukleinsäurefragmentes, welches für das gewünschte Protein kodiert, Ligation in einen Expressionsvektor, 15 Transformation in einen Wirt, Expression und Aufreinigung des Proteins (Fig. 25). In Übereinstimmung mit der vorliegenden Erfindung kann es vorkommen, daß Gene, insbesondere aus den Archaeobakterien, Inteine (*Proc. Natl. Acad. Sci. U S A* 1992 Jun 15;89(12):5577-5581, Intervening sequences in an Archaea DNA polymerase gene, Perler FB, Comb DG, Jack WE, Moran LS, 20 Qiang B, Kucera RB, Benner J, Slatko BE, Nwankwo DO, Hempstead SK, et al) enthalten können, welche zunächst entfernt werden können.

Die Identifizierung weiterer für den erfindungsgemäßen Komplex geeigneter Gene und/oder Proteine kann z.B. erfolgen durch Homologiesuchen in 25 Datenbanken, welche Genome aus Prokaryonten umfassen. Programme, die dafür geeignet sind, umfassen beispielsweise das Programm BLAST, BLASTP und FASTA, sind aber nicht darauf beschränkt. (Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new

generation of protein database search programs", Nucleic Acids Res. 25:3389-3402. W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448).

Die Identifizierung kann auch dadurch geschehen, daß DNA-Sonden
5 verwendet werden, um in z.B. gesamtgenomischen Banken aus Prokaryonten
oder Eukaryonten nach den entsprechenden Genen zu screenen. Die hierfür
nötigen experimentellen Verfahren finden sich in Maniatis et al. (*Molecular
Cloning (2nd edition, 3 Volume set): A laboratory Manual*, Cold Spring Harbor
Laboratory Press, N.Y.(1989)) oder in Ausubel et al (*Current Protocols in
10 Molecular Biology*, John Wiley and Sons (1988)).

Die Bereitstellung der gereinigten Nukleinsäure der Gene, die die
erfindungsgemäßen Komplexe aufbauenden Proteine kann z.B. über deren
Isolierung aus einer genomischen Bank des relevanten Organismus
15 geschehen oder durch synthetische DNA-Herstellung, jeweils gewünschtenfalls
kombiniert mit einer Amplifikation mittels PCR, unter Zuhilfenahme von
Primern, welche für den gewünschten Genabschnitt spezifisch sind. Übliche
Verfahren sind in Maniatis et al. (*Molecular Cloning: A laboratory Manual*, Cold
Spring Harbor Laboratory Press, N.Y. (1989)) beschrieben.

20

Die Gene der Proteine der erfindungsgemäßen *in vitro*-Komplexe können unter
Anwendung einer Vielzahl von Methoden kloniert werden und somit mittels
eines Expressionsvektors zur Proteinexpression in einem Wirtsorganismus zur
Verfügung gestellt werden. Gängige Verfahren sind in Maniatis et al. (*Molecular
25 Cloning: A laboratory Manual*, Cold Spring Harbor Laboratory Press,
N.Y.(1989)) beschrieben. Die Gene der Komplexe können hierbei z.B.
zunächst in einen 'high-copy' Vektor, z.B. pUC18, pBst oder pBR322 kloniert
werden und dann erst in einen prokaryontischen Expressionsvektor, z.B.
pTrc99, pQE30, pQE31 oder pQE32, umkloniert werden, oder aber direkt in
30 einen prokaryontischen oder viralen Expressionsvektor kloniert werden. Unter
Vektoren sind hierbei Nukleinsäuren zu verstehen, die in der Lage sind, ein

anderes Nukleinsäuremolekül in oder zwischen verschiedenen Organismen, bzw. genetischen Hintergründen zu transportieren. Sie haben in der Regel die Fähigkeit zur autonomen Replikation und/oder zur Expression (Expressionsvektoren) des operativ verbundenen Nukleinsäuremoleküls.

5 "Operativ verbunden" bedeutet hierin, daß das transportierte Nukleinsäuremolekül so mit dem Vektor verbunden ist, daß es unter der Transkriptions- und Translationskontrolle von Expressionskontrollsequenzen des Vektors steht und in einer Wirtszelle exprimiert werden kann. Bakterielle und virale Expressionssysteme, deren bevorzugte Verwendung, sowie eine

10 Auswahl an Vektorsystemen ist z.B. in *Gene Expression Technology*, (Meth. Enzymol., Vol 185, Goeddel, Ed., Academic Press, N.Y. (1990)) beschrieben. Geeignete Vektoren für die vorliegende Erfindung sollten eine unterschiedlich starke Expression der Proteine dadurch ermöglichen, daß sie einige oder alle der folgenden Eigenschaften besitzen: (1) Promotoren, oder

15 Transkriptionsinitiationsstellen, entweder unmittelbar neben dem Start des Proteins oder als Fusionsprotein, (2) Operatoren, verwendet werden können, um die Genexpression an oder aus zu schalten, (3) ribosomale Bindungsstellen für eine verbesserte Translation, und (4) Terminationsstellen für die Transkription oder Translation, die zu verbesserter Stabilität des Transkriptes

20 führen.

Expressionsvektoren, die mit eukaryontischen Zellen, bevorzugterweise mit Vertebratenzellen kompatibel sind, können auch Verwendung finden. Einige bekannte Vektoren sind pSVL und pKSV-10 (Pharmacia), pBPV-1/pML2d

25 (International Biotechnologies, Inc.), pAChLT-ABC (Pharmingen) und pTDT1 (ATCC 31255). Die Verwendung eines retroviralen Expressionsvektors ist auch möglich.

Weitere Gegenstände der vorliegenden Erfindung sind daher DNA-Sequenzen,

30 welche für die erfindungsgemäßen thermostabilen *in vitro*-Komplexe codieren, sowie entsprechende Vektoren, vorzugsweise Expressionsvektoren.

Die erfindungsgemäßen Vektoren enthalten mindestens ein Gen für das Gleitklammerprotein oder mindestens ein Gen für das Kopplungsprotein oder mindestens ein Gen für einen Gleitklammerlader 1 und/oder 2 oder mindestens
5 ein Gen für ein Elongationsprotein. In einer Ausführungsform enthalten die Vektoren gleichzeitig mehrere der verschiedenen vorstehend genannten Gene, beispielsweise die Gene für ein Elongationsprotein und ein Gleitklammerprotein.

10 Es ist dabei im Rahmen der Erfindung bevorzugt, wenn der Vektor neben den darin bereits enthaltenen DNA-Sequenzen noch geeignete Restriktionsschnittstellen und ggf. Polylinker zur Insertion weiterer DNA-Sequenzen enthält. Besonders bevorzugt ist es, wenn die räumliche
Anordnung von bereits enthaltener DNA-Sequenz und zusätzlicher
15 Insertionsstelle nach Expression zur Ausbildung eines Fusionsproteins führt.

Ebenfalls bevorzugt ist es, wenn der erfindungsgemäße Vektor Promotor- und/oder Operatorbereiche enthält, wobei es besonders bevorzugt ist, wenn derartige Promotor- und/oder Operatorbereiche induzierbar oder reprimierbar
20 sind. Dadurch wird eine Steuerung der Expression in Wirtszellen erheblich vereinfacht und kann besonders effizient gestaltet werden.

Derartige Promotor/Operatorbereiche können in einem Expressionsvektor auch mehrfach vorkommen, so daß ggf. eine unabhängige Expression mehrerer
25 DNA-Sequenzen unter Verwendung nur eines Expressionsvektors ermöglicht wird.

Ein weiterer Gegenstand der vorliegenden Erfindung ist eine Wirtszelle, enthaltend einen oder mehrere erfindungsgemäße(n) Vektor(en), wobei
30 dieser Wirtszelle unter geeigneten Bedingungen die Expression zu Proteinen

erfolgen kann. Geeignete Bedingungen schließen beispielsweise Anwesenheit eines Repressors, Induktors oder eines Derepressors ein.

Für die Transformation, Phageninfektion und Zellkultur existieren 5 Standardprotokolle in Maniatis et al. (*Molecular Cloning: A laboratory Manual*, Cold Spring Harbor Laboratory Press, N.Y.). Aus der Vielzahl der vorhandenen *E. coli*-Stämme, die zur Transformation geeignet sind, sind bevorzugt JM101 (ATCC No. 33876), XL1 (Stratagene), RRI (ATCC No. 31343)M15[prep4] (QIAGEN) und BL21 (Pharmacia). Proteinexpression kann z.B. auch 10 geschehen unter Zuhilfenahme des *E. coli* Stammes INV α F' (Invitrogen).

Die Transformanten werden unter den geeigneten Wachstumsbedingungen des Wirtsstammes kultiviert. So werden die meisten *E. coli*-Stämme z.B. in LB-Medium bei 30°C bis 42°C bis zur logarithmischen oder stationären Wachstumsphase kultiviert. Die Proteine können aus einer transformierten 15 Kultur gereinigt werden, wobei dies entweder aus einem Zellpellet, nach Zentrifugation, oder aber aus der Kulturflüssigkeit geschehen kann. Sofern die Proteine aus dem Zellpellet gereinigt werden, werden die Zellen in einem geeigneten Puffer resuspendiert und mittels Ultraschallbehandlung, enzymatischer Behandlung oder Einfrieren und Auftauen aufgebrochen. Sofern 20 die Reinigung aus der Kultursuspension geschieht, sei es ohne oder mit einem Fusionsprotein, wird der Überstand von den Zellen mittels bekannter Verfahren, wie Zentrifugation abgetrennt.

Die Trennung und Reinigung der Proteine der erfindungsgemäßen Komplexe 25 entweder aus dem Überstand der Kulturlösung oder aus dem Zellextrakt kann durch bekannte Separations- oder Reinigungsverfahren geschehen. Diese Methoden sind z.B. solche, die die Löslichkeiten betreffen, wie Salzfällungen und Lösungsmittelfällungen, Methoden, die sich die unterschiedlichen Molekulargewichte zunutze machen, wie Dialyse, Ultrafiltration, Gelfiltration, 30 und SDS-Polyacrylamidgelelektrophorese, Methoden die sich die unterschiedlichen Ladungen zunutze machen, wie

Ionenaustauschchromatographie, Methoden die sich die unterschiedliche Hydrophobie zunutze machen, wie *reverse-phase* HPLC (High Performance Liquid Chromatography), Methoden, die sich bestimmte Affinitäten zunutze machen, wie Affinitätschromatographie (Beispiel 6, 7 Fig. 24 und Fig. 25), und 5 Methoden, die sich Unterschiede im isoelektrischen Punkt zunutze machen, wie isoelektrische Fokussierung. Es ist auch vorstellbar, daß Zellextrakte, entweder aus dem Organismus, welcher das Gen des akzessorischen Komplexes trägt, gemacht werden können, welcher die erfindungsgemäße Aufgabe erfüllt, oder aus dem rekombinanten Wirtsorganismus, z.B. *E. coli*. Mit 10 diesen Extraktionen könnte man andere Reinigungsschritte vermeiden.

Die oben beschriebenen Verfahren können in einer Vielzahl von Kombinationen eingesetzt werden, um die Proteine des *in vitro*-Komplexes bereitzustellen.

15

Elongation von Nukleinsäuren

Der erfindungsgemäße thermostabile *in vitro*-Komplexe kann zur Elongation von Nukleinsäuren verwendet werden, z. B. zur Polymerase-Ketten-Reaktion, 20 (Beispiel 3, 4, sowie Fig. 21; Fig. 22) DNA-Sequenzierung, zur Markierung von Nukleinsäuren und anderen Reaktionen, die die *in vitro* Synthese von Nukleinsäuren beinhalten.

Ein weiterer Gegenstand der vorliegenden Erfindung ist daher ein Verfahren 25 zur Template-abhängigen Elongation von Nukleinsäuren, wobei die Nukleinsäure nötigenfalls denaturiert, mit mindestens einem Primer unter Hybridisierungsbedingungen versehen wird, wobei der Primer hinreichend komplementär zu einem flankierenden Bereich einer gewünschten Nukleinsäuresequenz des Template-Strangs ist und mit Hilfe einer Polymerase 30 in Anwesenheit von Nukleotiden eine Primer-Elongation erfolgt, wobei als

Polymerase ein erfindungsgemäßer thermostabiler *in vitro*-Komplex eingesetzt wird.

Verfahren zur Template-abhängigen Elongation von Nukleinsäuren, bei denen
5 die Elongation ausgehend von einem Primer erfolgt, der an die Template-Nukleinsäure anhybridisiert wurde und ein freies 3'-OH-Ende für die Elongation zur Verfügung stellt, sind dem Fachmann bekannt. Zur Amplifikation wird insbesondere eine PCR-Polymerasekettenreaktion durchgeführt.

10 Reverse Transkription

Es ist auch die Anwendung des erfindungsgemäßen thermostabilen *in-vitro*-Komplexes in der reversen Transkription möglich, wobei entweder der erfindungsgemäße Komplex selbst Reverse-Transkriptase-Aktivität besitzt,
15 Reverse-Transkriptase-Aktivität aufweist, unabhängig davon ob der thermostabile *in vitro*-Komplex selbst eine Reverse-Transkriptase-Aktivität hat.

Zur erfindungsgemäß bevorzugten reversen Transkription von RNA in DNA wird ebenfalls ein erfindungsgemäßer thermostabiler *in vitro*-Komplex
20 eingesetzt, wobei dessen Elongationsprotein selbst eine Reverse Transkriptase-Aktivität aufweist. Diese Reverse Transkriptase-Aktivität kann die einzige Polymeraseaktivität des Elongationsproteins sein, kann aber auch zusätzlich zu einer vorhandenen 5'-3'-DNA-Polymeraseaktivität vorliegen. Eine erfindungsgemäß bevorzugte Ausführungsform des *in vitro*-Komplex umfaßt
25 das Elongationsprotein aus dem Organismus *Carboxydothemus hydrogenoformans*, wie offenbart in EP-A 0 834 569.

Sequenzierung

Eine weitere bevorzugte Verwendung des erfindungsgemäßen *in vitro*-Komplexes ist die Sequenzierung von Nukleinsäuren gemäß der Methode von Sanger. Ausgehend von mindestens einem Primer, der zu einem Teil der zu sequenzierenden Nukleinsäure hinreichend komplementär ist, wird eine
5 Template-abhängige Elongation vorgenommen. Im Falle der Sequenzierung von RNA ist es nötig, eine reverse Transkription durchzuführen. Als Deoxynukleotide oder Dideoxynukleotide werden im Rahmen dieser bevorzugten Ausführungsform auch die oben beschriebenen jeweiligen
10 erfindungsgemäßen Verfahren zur Elongation von Nukleinsäuren bevorzugt, daß die gebildeten Nukleinsäuren markiert werden. Hierzu ist es möglich, markierte Primer und/oder markierte Deoxynukleotide und/oder markierte Dideoxynukleotide und/oder markierte Ribonukleotide oder jeweils entsprechende Derivate, wie sie oben bereits beispielsweise beschrieben sind,
15 einzusetzen.

Markierung von Nukleinsäuren

Ein weiterer Gegenstand der vorliegenden Erfindung ist ein Verfahren zur
20 Markierung von Nukleinsäuren, z.B. durch Einfügung einzelner Brüche in Phosphodiesterbindungen der Nukleinsäurekette und Ersatz eines Nukleotids an den Bruchstellen durch ein markiertes Nukleotid mit Hilfe einer Polymerase, wobei als Polymerase ein erfindungsgemäßer thermostabiler *in vitro*-Komplex eingesetzt wird.

25

Ein bevorzugtes Verfahren stellt das allgemein als Nick-Translation bezeichnete Verfahren dar, das eine einfache Markierung von Nukleinsäuren ermöglicht. Alle oben bereits beschriebenen markierten Ribonukleotide oder Deoxyribonukleotide oder Derivate davon sind hierfür geeignet, solange die
30 Polymerase sie als Substrat akzeptiert.

Markierung im vorstehenden Sinne ist auch eine Markierung, die im Rahmen einer PCR-Reaktion erfolgt, wobei in diesem Fall markierte Nukleotide oder Derivate davon in die Nukleinsäuresequenz eingebaut werden.

5 Gegenstand der vorliegenden Erfindung ist ebenfalls ein Kit zur Elongation und/oder Amplifikation und/oder Reversen Transkription und/oder Sequenzierung von Nukleinsäuren, wobei dieser Kit ein oder mehrere Behälter umfassen kann.

10 Der Kit selbst umfaßt

- a) einen erfindungsgemäßen thermostabilen *in vitro*-Komplex oder
- b) einen thermostabilen *in vitro*-Komplex und ggf. separat davon ein Polymeraseaktivität aufweisendes Elongationsprotein sowie ggf. eine oder mehrere der Komponenten, die ausgewählt sind aus der Gruppe, die Primer, Puffersubstanzen, Nukleotide, ATP, andere Kofaktoren und Pyrophosphatase umfaßt.

15 Dabei ist es möglich, daß der erfindungsgemäße thermostabile *in vitro*-Komplex in besagtem Kit in Form seiner Einzelbestandteile, das heißt als 20 thermostabiles Gleitklammerprotein sowie thermostabiles Elongationsprotein vorliegt, getrennt oder vereint in einem Behältnis, und sich als solches erst zu einem späteren Zeitpunkt ausbildet.

Insbesondere ist Gegenstand der vorliegenden Erfindung ein Kit zur 25 Elongation, Amplifikation, Reversen Transkription, Markierung bzw. Sequenzierung von Nukleinsäuren, wobei zusätzlich Deoxynukleotide bzw. deren Derivate enthaltend sind.

Gegebenenfalls können in dem erfindungsgemäßen Kit auch Ribonukleotide 30 oder Derivate davon umfaßt sein, nämlich dann, wenn ein Elongationsprotein eingesetzt wird, das Ribonukleotide als Substrat akzeptiert.

Eine bevorzugte Ausführungsform des erfindungsgemäßen Kits ist ein Kit zur Sequenzierung von Nukleinsäuren, wobei zusätzlich zu Deoxynukleotiden bzw. deren Derivaten, Dideoxynukleotide bzw. deren Derivate zur 5 Kettentermination umfaßt sind.

Desweiteren ist insbesondere Gegenstand der vorliegenden Erfindung ein Kit zur reversen Transkription von Nukleinsäuren, wobei entweder der erfindungsgemäße Komplex selbst reverse Transkriptase-Aktivität besitzt, 10 oder aber zusätzlich ein geeignetes Enzym anwesend ist, das reverse Transkriptase-Aktivität besitzt, wobei Deoxynukleotide bzw. deren Derivaten im Reaktionsgemisch enthalten sein können.

In einer weiteren besonders bevorzugten Ausführungsform enthält der Kit 15 Primer und/oder Deoxynukleotide und/oder Dideoxynukleotide und/oder Ribonukleotide und/oder deren jeweilige Derivate in markierter Form.

Insbesondere für die Sequenzierung von Nukleinsäuren ist es nötig, eine Markierung einzufügen. Geeignete Markierungen sind weiter oben bereits in 20 beispielhafter Form beschrieben. Geeignete Markierungsmittel sind in einer bevorzugten Ausführungsform des erfindungsgemäßen Kits umfaßt.

Im Rahmen der vorliegenden Erfindung ist es weiterhin möglich, daß der Kit (hierin auch als Reagenzienkit bezeichnet) zur Markierung von 25 Nukleinsäuren eingesetzt wird. In diesem Fall enthält der Reagenzienkit die Bestandteile a) oder b) und markierte Nukleotide, wobei Puffersubstanzen, ATP oder andere Kofaktoren und/oder Pyrophosphatase ebenfalls anwesend sein können.

Insbesondere ist ein Kit bevorzugt, der einen geeigneten Puffer umfaßt, wie oben beschrieben. Bevorzugt ist ebenfalls, daß der erfindungsgemäße Kit zusätzlich eine Pyrophosphatase, ATP und/oder andere Kofaktoren umfaßt.

- 5 Gegenstand der vorliegenden Erfindung ist desweiteren die Verwendung eines thermostabilen Gleitklammerproteins in *in vitro* Verfahren zur Elongation, Amplifikation, Markierung bzw. Sequenzierung oder reversen Transkription von Nukleinsäuren.
- 10 Der erfindungsgemäße thermostabile *in vitro*-Komplex kann zu Zwecken der Sequenzierung, Amplifikation, reversen Transkription und dergleichen sowohl mit kurzen als auch mit langen Nukleinsäurefragmenten verwendet werden, wobei insgesamt eine verringerte Fehlerhäufigkeit (Nukleotidfehlinkorporation) gegenüber der Verwendung thermostabiler
- 15 Elongationsproteine alleine erreicht wird.

In einer bevorzugten Alternative wird der thermostabile *in vitro*-Komplex zu Zwecken der Sequenzierung, Amplifikation und reversen Transkription von langen Nukleinsäurefragmenten verwendet.

20

In einer weiteren bevorzugten Alternative wird der thermostabile *in vitro*-Komplex zu Zwecken der Sequenzierung, Amplifikation und reversen Transkription von solchen Nukleinsäurefragmenten verwendet, die Sekundärstruktur ausbilden.

25

Die folgenden Beispiele sollen in Verbindung mit den Figuren die Erfindung näher erläutern und Vorteile veranschaulichen:

Figuren:

30

Im folgenden werden häufig Sequenznamen verwendet, unter denen die Protein- oder Nukleinsäuresequenzen in der Genbank und der EMBL Datenbank stehen. Dabei zeigt:

- 5 Fig. 1 tabellarisch Proteinsequenznamen des erfindungsgemäßen *in vitro*-Komplexes, sowie Werte aus paarweisen Alignments und multiplen Alignments zwischen Archaeobakterien und zwischen Archaeobakterien und den entsprechenden humanen Genen;
- 10 Fig. 2 eine Tabelle ähnlich der von Fig. 1, jedoch beschränkt auf Gleitklammerprotein und Elongationsprotein von *E. coli* und *A. aeolicus*;
- Fig. 3 eine schematische Darstellung des erfindungsgemäßen thermostabilen *in vitro*-Komplexes;
- 15 Fig. 4 Alignments zweier konservierter Regionen des Gleitklammerproteins;
- Fig. 5 ein Alignment einer konservierten Region des
- 20 Kopplungsproteins;
- Fig. 6 ein Alignment einer konservierten Region des Gleitklammerladers;
- 25 Fig. 7 ein Alignment einer konservierten Region des Gleitklammerladers 2;
- Fig. 8 ein Alignment einer konservierten Region des Elongationsproteins 1;

Fig. 9 ein Alignment einer konservierten Region des Elongationsproteins 2;

Fig. 10 ein Alignment einer konservierten Region des 5 Elongationsproteins aus Eubakterien;

Fig. 11 ein Alignment einer konservierten Region der Gleitkammer aus Eubakterien;

10 Fig. 12 bis 19 multiple Alignments verschiedener Proteinsequenzen zur Generierung von Hidden Markov Modellen (HMM);

Fig. 20 eine chromatographische Analyse einer rekombinanten Gleitkammer (*Archaeoglobus fulgidus* PCNA (AF 0335));

15

Fig. 21 und 22 das Ergebnis einer PCR unter Verwendung eines Elongationsproteins, wie es Bestandteil eines erfindungsgemäßen thermostabilen *in vitro*-Komplexes ist;

20 Fig. 23 den Vergleich der Aktivität eines Elongationsproteins mit und ohne Gleitkammerprotein;

Fig. 24 das Resultat der Reinigung eines Gleitkammerproteins;

25 Fig. 25 die Expression und Reinigung der *Archaeoglobus fulgidus*-DNA-Polymerase;

Fig. 26 die Ergebnisse der Verwendung eines erfindungsgemäßen *in vitro*-Komplexes in der PCR;

30

Fig. 27 die Ergebnisse eines Y2H Experimentes;

Fig. 28 das PCR-Amplifikationsergebnis des humanen Kollagen-Gens unter Verwendung des erfindungsgemäßen thermostabilen *in vitro*-Komplex

5 Fig. 29 eine tabellarische Übersicht derjenigen Gene, wie sie den in Fig. 1 angegebenen Proteinsequenznummern entsprechen und aus den jeweiligen Datenbanken erhalten werden können; und

Fig. 30 eine tabellarische Darstellung der Hintergrundinformation zu
10 den verschiedenen Datenbanken in englischer Sprache, aus denen die hierin angegebenen Nukleinsäuresequenz- und Aminosäuresequenzdaten erhalten werden können.

Detaillierte Beschreibung der Figuren

15

Fig. 1:

Figur 1 listet tabellarisch Proteinsequenznamen des erfindungsgemäßen thermostabilen *in vitro*-Komplexes auf sowie Werte aus paarweisen
20 Alignments und multiplen Alignments zwischen Archaeobakterien und zwischen Archaeobakterien und den entsprechenden humanen Genen. Hierbei bedeutet die Annotation „¹“ die Prozentidentität (%) zu dem entsprechenden humanen Gen, berechnet aus dem paarweisen Alignment (siehe Figuren) mit BLASTP 2.0.4 [Feb-24-1998] und die Annotation „²“
25 Prozentidentität zu dem entsprechenden Gen von *Archaeoglobus fulgidus*, berechnet aus dem paarweisen Alignment (siehe Figuren) mit BLASTP 2.0.4 [Feb-24-1998] und die Annotation „³“ Prozentidentität zu dem entsprechenden humanen Gen, berechnet aus dem paarweisen Alignment mit FASTA 3.1t02 [March, 1998]⁵. Die Verfahren werden näher beschrieben
30 in: Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped

BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402 und W.R. Pearson & D.J. Lipman Proc. Natl. Acad. Sci. (1988) 85:2444-2448. Die Fig. 1 zeigt im Falle des Gleitklammerladers 1, des Gleitklammerlader 2, des Gleitklammer, des Kopplungsprotein und des Elongationsprotein 1 die Sequenznamen aus den Datenbanken und ebenso deren SEQ ID Nummer, wobei die Werte in den Klammern jeweils Prozent Identität je Anzahl Aminosäuren darstellen. Im Falle des Elongationsprotein 2 beziehen sich die Werte auf Prozent Sequenzidentität zu der *Archaeoglobus fulgidus*- Sequenz. Auch hier sind die Sequenznamen aus den Datenbanken und ebenso deren "SEQ ID No." gezeigt.

Fig. 2:

Fig. 2 zeigt Proteinsequenzen, paarweise Alignments und multiple Alignments aus Eubakterien des Replikationsapparates. Wobei die Annotation ¹ %Identität zu dem entsprechenden Gen aus *E. coli* bedeutet, berechnet aus dem paarweisen Alignment (siehe Anhang) mit BLASTP 2.0.4 [Feb-24-1998]#. Das Verfahren wird in: Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402, näher beschrieben.

Fig. 3:

Fig. 3 zeigt eine Skizze einer möglichen Form des erfindungsgemäßen thermostabilen *in vitro*-Komplexes, wobei die Gleitkammer über ein Kopplungsprotein an das Elongationsprotein bindet.

Fig. 4:

Fig. 4 zeigt Alignments zweier konservierter Regionen des Gleitklammerproteins, sowie die daraus hergeleiteten Konsensussequenzen. Folgende Gene sind gezeigt: PCNA human (aus SEQ ID NO: 11) und die entsprechenden Sequenzen aus *Archaeoglobus fulgidus* (aus SEQ ID NO: 12), aus *Methanococcus janashii* (aus SEQ ID NO: 13), aus *Pyrococcus horikoshii* (aus SEQ ID NO: 14) und aus *Methanococcus thermoautotrophicus* (aus SEQ ID NO: 15).

10

Fig. 5:

Fig. 5 zeigt ein Alignment einer konservierten Region des Kopplungsproteins sowie die daraus hergeleiteten Konsensussequenzen. Folgende Gene sind gezeigt: PfuORF2, DPD2_HUMAN, AF1790 und MJ0702. Die SEQ ID Nummern können Fig. 1 entnommen werden.

Fig. 6:

Fig. 6 zeigt ein Alignment einer konservierten Region der koppelnden Untereinheit sowie die daraus hergeleiteten Konsensussequenzen. Folgende Gene sind gezeigt: AC11_HUMAN, AF2060, MTH0241, PHBN012 und MJ1422. Die SEQ ID Nummern können Fig. 1 entnommen werden.

Fig. 7:

Fig. 7 zeigt ein Alignment einer konservierten Region des Gleitklammerlader 2 sowie die daraus hergeleiteten Konsensussequenzen. Folgende Gene sind gezeigt: AC15_HUMAN, MJ0884, AF1195, MTH0240 und MTH0240. Die SEQ ID Nummern können Fig. 1 entnommen werden.

Fig. 8:

Fig. 8 zeigt ein Alignment einer konservierten Region des
5 Elongationsproteins 1, sowie die daraus hergeleiteten Konsensussequenzen.
Folgende Gene sind gezeigt: DPOD_HUMAN, MJ0885 , MTH1208,
PHBT047 und DPOL_ARCFU. Die SEQ ID Nummern können Fig. 1
entnommen werden.

10 Fig. 9:

Fig. 9 zeigt ein Alignment einer konservierten Region des
Elongationsproteins 2, sowie die daraus hergeleiteten Konsensussequenzen.
Folgende Gene sind gezeigt: AF1722, MJ1630 , PfuORF3, MTH1536 und
15 PHBN021. Die SEQ ID Nummern können Fig. 1 entnommen werden.

Fig. 10:

Fig. 10 zeigt ein Alignment einer konservierten Region des
20 Elongationsproteins aus Eubakterien sowie die daraus hergeleiteten
Konsensussequenzen. Folgende Gene sind gezeigt: DP3A_ECOLI: DNA Pol
III, alpha subunit, *Escherichia coli*, BB0579: DNA Pol III, alpha subunit,
Borrelia burgdorferi, DP3A_HELPY: DNA Pol III, alpha subunit, *Helicobacter*
pylori AA50: *Aquifex aeolicus*, section 50 und DP3A_SALTY: DNA Pol III,
25 alpha subunit, *Salmonella typhimurium*).

Fig. 11:

Fig. 11 zeigt ein Alignment einer konservierten Region der Gleitklammer aus
30 Eubakterien sowie die daraus hergeleiteten Konsensussequenzen. Folgende

Gene sind gezeigt: AAPOL3B, DP3B_ECOLI, S.TYPHIM, DP3B_PROMI, DP3B_PSEPU und DP3B_STRCO (AAPOL3B: *Aquifex aeolicus* sektion 93: DP3B_ECOLI : DNA Pol III, beta chain, *Escherichia coli* S.TYPHIM: DNA Pol III, beta chain, *Salmonella typhimurium* P3B_PROMI : DNA Pol III, beta chain, *Proteus mirabilis* DP3B_PSEPU: DNA Pol III, beta chain, *Pseudomonas putida* DP3B_STRCO: DNA Pol III, beta chain, *Streptomyces coelicolor*).

Fig. 12:

10

Fig. 12 zeigt ein multiples Alignment der Gleitklammerprotein-Sequenzen zur Generierung des HMM.

Fig. 13:

15

Fig. 13 zeigt ein multiples Alignment der eubakteriellen Gleitklammerprotein-Sequenzen zur Generierung des HMM (AAPOL3B: *Aquifex Aeolicus* sektion 93: DP3B_ECOLI : DNA Pol III, beta chain, *Escherichia coli* S.TYPHIM: DNA Pol III, beta chain, *Salmonella typhimurium* P3B_PROMI : DNA Pol III, beta chain, *Proteus mirabilis* DP3B_PSEPU: DNA Pol III, beta chain, *Pseudomonas putida* DP3B_STRCO: DNA Pol III, beta chain, *Streptomyces coelicolor*).

25 Fig. 14:

Fig. 14 zeigt ein multiples Alignment der Gleitklammerlader 1-Proteinsequenzen zur Generierung des HMM.

Fig. 15:

Fig. 15 zeigt ein multiples Alignment der Gleitklammerlader 2-Proteinsequenzen zur Generierung des HMM.

5

Fig. 16:

Fig. 16 zeigt ein multiples Alignment der Proteinsequenzen der Kopplungsproteine zur Generierung des HMM.

10

Fig. 17:

Fig. 17 zeigt ein multiples Alignment der Sequenzen der Elongationsproteine 1 zur Generierung des Hidden Markov Models.

15

Fig. 18:

Fig. 18 zeigt ein multiples Alignment der Sequenzen der Elongationsproteine 2 zur Generierung des Hidden Markov Models.

20

Fig. 19:

Fig. 19 zeigt ein multiples Alignment der Sequenzen der eubakteriellen Elongationsproteine zur Generierung des HMM. Folgende Gene sind gezeigt:

25 DP3A_ECOLI: DNA Pol III, alpha subunit, *Escherichia coli*, BB0579: DNA Pol III, alpha subunit, *Borrelia burgdorferi*, DP3A_HELPY: DNA Pol III, alpha subunit, *Helicobacter pylori* AA50: *Aquifex aeolicus*, section 50 und DP3A_SALTY: DNA Pol III, alpha subunit, *Salmonella typhimurium*).

Fig. 20: (Beispiel 2)

Fig. 20 zeigt eine chromatographische Analyse von rekombinatem *Archaeoglobus fulgidus* PCNA (AF 0335):

Rekombinantes *Archaeoglobus fulgidus*-PCNA (AF 0335) liegt unter nativen Bedingungen als Trimer vor. Fig. 20 A zeigt Proteine mit His-Tag (Histidin-Tag) in den Fraktionen 15 (Spur 1), 17 (Spur 2), 19 (Spur 3), 21 (Spur 4), 23 (Spur 5), 25 (Spur 6) der ohne Harnstoff durchgeführten Chromatographie. Fig. 20B zeigt Proteine mit His-Tag in den Fraktionen 10 (Spur 1), 11 (Spur 2), 12 (Spur 3), 13 (Spur 4), 14 (Spur 5), 15 (Spur 6), 16 (Spur 7), 17 (Spur 8), 10 der in Anwesenheit von Harnstoff durchgeführten denaturierenden Chromatographie.

Fig. 21: (Beispiel 3)

Fig. 21 zeigt das Ergebnis einer PCR unter Verwendung eines Elongationsproteins, wie es Bestandteil eines erfindungsgemäßen thermostabilen *in vitro*-Komplexes ist:

Jeweils 1 µl (Spur 4), 2,5 µl (Spur 5) und 5 µl (Spur 6) *Pyrococcus horikoshii* DNA-Polymerase (PH1947; Grobextrakt s. Fig. 25) wurden zum Aktivitätsvergleich mit 1 Einheit Taq Polymerase (Spur 2) und 1 µl *Archaeoglobus fulgidus* DNA-Polymerase (AF 0497)-Grobextrakt (s. Fig. 25) (Spur 3) individuell in Standard PCR Reaktionen eingesetzt.; Spur 1 zeigt einen DNA-Größenmarker (New England Biolabs; Mischung aus 1 kb DNA ladder und 100 bp DNA ladder).

Fig. 22: (Beispiel 4)

Fig. 22 zeigt das Ergebnis einer PCR unter Verwendung eines Elongationsproteins, wie es Bestandteil eines erfindungsgemäßen thermostabilen *in vitro*-Komplexes ist:

Proben der PCR mit *Archaeoglobus fulgidus* DNA-Polymerase AF 0497 wurde nach unterschiedlichen Zyklenzahlen (Z) entnommen und auf einem

1% Agarosegel in 1 x TAE Puffer (40 mM Tris-Acetat; 20 mM Natriumacetat; 10 mM EDTA ; pH 7,2) mit 10V/cm aufgetrennt. Spur 2: 16Z; Spur 3: 21Z; Spur 4: 26Z; Spur 5: 28Z; Spur 6: 30Z; Spur 7: 32Z; Spur 8: 34Z; Spur 9: 36Z; Spur 10: 38Z; Spur 11: 40Z; Spur 1 zeigt einen DNA Größenmarker 5 (New England Biolabs; Mischung aus 1 kb DNA ladder und 100 bp DNA ladder). Der obere Abschnitt zeigt Reaktionsprodukte der Taq Polymerase; der untere Abschnitt zeigt Reaktionsprodukte der *Archaeoglobus fulgidus* DNA-Polymerase AF 0497.

10 Fig. 23: (Beispiel 5)

Fig. 23 zeigt einen Vergleich der Aktivität eines Elongationsproteins mit und ohne Gleitklammerprotein. Probe 1 repräsentiert einen enzymfreien Ansatz. Proben 2-12 enthielten außerdem je 3 µl einer 1:1000 Verdünnung einer 15 Fraktion von rekombinanter *Archaeoglobus fulgidus*-DNA-Polymerase (Ausgangskonzentration 7,5 µg/µl). Proben 3-7 sowie 8-12 außerdem 0,5; 1; 2; 4 und 8 µl einer Fraktion von rekombinantem Gleitklammerprotein aus *Archaeoglobus fulgidus* PCNA. Intensitätsauswertung mit AIDA: Intensität Hintergrund Spur 1: 46,4; Spur 2= 258,5; Spur 3= 164,4; Spur 4= 122,8; 20 Spur 5= 162,1 ; Spur 6= 297,4; Spur 7=359,5

Fig. 24: (Beispiel 6)

Fig. 24 zeigt das Ergebnis der Reinigung eines Gleitklammerproteins: Spur 1 zeigt einen Molekulargewichtsstandard ((BIO RAD Cat Nr 161-0317). Für 25 Spur 2 wurden 500 µl Bakterien direkt vor der Induktion sedimentiert, gemäß der Anweisung des Herstellers zum Lauf von NuPage Gelen (NOVEX; Fig. 25) behandelt und aufgetragen. Spur 3 zeigt die gleiche Menge Bakterien 16 Stunden nach der Elution. Spuren 4 und 5 zeigen je 8 µl der beiden Eluate der Ni-NTa Agarosesäule nach der Dialyse. Durch Reinigung über Ni-NTA-

Agarose (Qiagen) werden hochreine Fraktionen des Gleitklammerproteins des Organismus *Archaeoglobus fulgidus* erhalten (siehe Spuren 4 und 5).

Fig. 25: (Beispiel 7)

5 Fig. 25 zeigt die Expression und Reinigung der *Archaeoglobus fulgidus* DNA-Polymerase (siehe auch Beispiel 7):
Spur 1 zeigt einen Molekulargewichtsstandard ((BIO RAD Cat Nr 161-0317).
Für Spur 2 wurden 500µl Bakterien direkt vor der Induktion sedimentiert,
gemäß der Anweisung zum Lauf von NuPage Gelen behandelt und
10 aufgetragen. Spur 3 zeigt die gleiche Menge Bakterien 16 Stunden nach der
Elution. Spuren 4 und 5 zeigen je 8µl der beiden Eluate der Ni-Nta
Agarosesäule nach der Dialyse. Spur 6 zeigt 8 ml eines dialysierten
Grobextraktes. Spuren 4 und 5 zeigen hochreine Fraktionen der
Archaeoglobus fulgidus DNA-Polymerase, die durch Reinigung über Ni-NTA-
15 Agarose werden erhalten.

Fig. 26: (Beispiel 8)

In Fig. 26 sind die Ergebnisse der Verwendung eines erfindungsgemäßen *in vitro*-Komplexes in der PCR gezeigt. Spur eins zeigt eine PCR Reaktion
20 ohne Verwendung einer Gleitklammer, wohingegen in der Spur 2 das
Ergebnis einer PCR Reaktion unter Verwendung eines Gleitklammerproteins
gezeigt ist.

Fig. 27: (Beispiel 9)

25

Fig. 27 zeigt die Ergebnisse eines „Yeast-Two-Hybrid“-Experiments, hierin
als Y2H- Experiment bezeichnet, wobei die Reihe A mit Zellen bestückt ist,
die den leeren pGAD424 Vektor (Clontech, Palo Alto, USA) tragen, sodaß
die Transkriptions-Aktivierungsdomäne exprimiert wird, die Reihe B mit
30 Zellen bestückt ist, die den pGAD424 Vektor tragen, von dem das

- Sacharomyces cerevesiae*-Gen CDC48 als Fusionsprotein mit der Transkriptions-Aktivierungsdomäne exprimiert wird, die Reihe C mit Zellen bestückt ist, die den pGAD424 Vektor tragen, von dem das Gleitklammergen aus *Archaeoglobus fulgidus* als Fusionsprotein mit der Transkriptions-
- 5 Aktivierungsdomäne exprimiert wird, die Reihe D nicht mit Zellen bestückt ist und die Reihe E mit Zellen bestückt ist, die den pGAD424 Vektor tragen, von dem das Elongationsproteingen aus *Archaeoglobus fulgidus* als Fusionsprotein mit der Transkriptions-Aktivierungsdomäne exprimiert wird.
- 10 Die Spalte 1 ist mit Zellen bestückt, die den leeren pGBT9 Vektor (Clontech, Palo Alto, USA) tragen, sodaß die DNA-Bindedomäne exprimiert wird, die Spalte 2 ist bestückt mit Zellen, die den pGBT9 Vektor tragen, von dem das *Saccharomyces cerevisiae*-Gen UFD3 als Fusionsprotein mit der DNA-Bindedomäne exprimiert wird, die Spalte 3 ist bestückt mit Zellen, die den
- 15 pGBT9-Vektor tragen, von dem das Gleitklammerprotein aus *Archaeoglobus fulgidus* als Fusionsprotein mit der DNA-Bindedomäne exprimiert wird, die Spalte 4 ist bestückt mit Zellen, die den pGBT9 Vektor tragen, von dem das Kopplungsprotein aus *Archaeoglobus fulgidus* als Fusionsprotein mit der DNA-Bindedomäne exprimiert wird und die Spalte 5 ist bestückt mit Zellen,
- 20 die den pGBT9 Vektor tragen, von dem das Elongationsprotein aus *Archaeoglobus fulgidus* als Fusionsprotein mit der DNA-Bindedomäne exprimiert wird.

Fig. 28: (Beispiel 10)

- 25 In Fig. 28 ist das PCR-Amplifikationsergebnis des humanen Kollagen-Gens jeweils unter Verwendung des erfindungsgemäßen thermostabilen *in vitro*-Komplexes und ohne diesen dargestellt. Das erwartete Amplifikat hat in beiden Fällen eine Größe von etwa 1 Kb. Spur 1 zeigt einen Molekulargewichtsmarker, Spur 2 zeigt das Ergebnis der Amplifikation unter
- 30 Verwendung eines erfindungsgemäßen Elongationsproteins ohne

Gleitklammer und Spur 3 zeigt das Ergebnis der Amplifikation unter Verwendung des erfindungsgemäßen thermostabilen *in vitro*-Komplexes.

Beispiele:

5

Die Erfindung wird durch die folgenden Beispiele näher beschrieben, ist aber nicht auf diese beschränkt.

Beispiel 1:

10

Die DNA wird mittels der bekannten Verfahren aus dem Organismus *Archaeoglobus fulgidus* (DSM No. 4304) gereinigt. Die Aufzucht der Organismen erfolgte hierbei durch die DSM (Deutsche Sammlung für Mikroorganismen). Um die entsprechenden Gene (Gleitklammerlader 1/2, Gleitklammer, Elongationsproteine 1/2, koppelnde Untereinheit) in die Expressionsvektoren pTrc99 und pQE30 zu klonieren, werden für jedes Gen Primer entwickelt, die den vollständigen offenen Leserahmen samt Stopcodon umspannen. Nur für die Klonierung in pTRC99 ist in den Primer-Sequenzen ebenfalls das Startcodon enthalten. Die entsprechenden Primer zur Klonierung in pQE30 enthalten die dem Startcodon unmittelbar folgenden Nukleotide als erste genspezifische Sequenzen. Hierbei werden den Primern jeweils Restriktionsenden hinzugefügt, die die gerichtete Klonierung in den Expressionsvektor erleichtern. Unter Verwendung von etwa 200 ng gesamtgenomischer DNA werden mit den entsprechenden Annealing-Temperaturen PCR-Reaktionen gefahren (etwa 35 Zyklen) und die daraus resultierenden Produkte gereinigt. Nach der Reinigung werden die Produkte mit Restriktionsenzymen behandelt und über ein Agarosegel gereinigt, um zur Ligation bereit zu stehen. Der Expressionsvektor wird mittels Restriktionsenzymen so linearisiert, gereinigt und verdünnt, daß er zur Ligation mit den Amplifikaten der Gene des erfindungsgemäßen *in vitro*-

Komplexes aus der obigen PCR bereit ist. Die Ligation wird angesetzt und nach Inkubation ein Aliquote in einen der *E. coli* Stämme INValphaF' (Invitrogen) oder XL1 blue oder M15 [prep4] transformiert. Von jedem Gen werden mindestens 3 positive Kolonien gepickt, Plasmid-DNA präpariert und die Inserts auf Vollständigkeit und Richtigkeit mittels DNA-Sequenzierung überprüft. Korrekte Klone werden ausgesucht und erneut zur Vereinzelung auf Agarplatten (Ampicillin; Ampicillin und Kanamycin) ausgebracht. Kolonien werden gepickt und Übernachtskulturen angesetzt. Ein Aliquot (500 µl) der Übernachtskultur wird in eine ein bis fünf Liter Kultur von LB (Ampicillin: 80 mg/l bzw. zusätzlich 25 µg/ml Kanamycin für M15 [prep4] Stämme) gegeben. Die Kulturen wachsen bis zu einer OD₆₆₀ von 0.8 bei 37°C. Nun wird zur Induktion IPTG zugegeben (125 mg/l). Diese Kulturen wachsen nun weitere 4-21 Stunden. Die Kulturen werden zentrifugiert und ausgehend von dem Vektor pQE30 exprimierte rekombinante Proteine gemäß protocol 8 und 11 aus „The QIAexpressionist“ (Dritte Auflage, QIAGEN) extrahiert und gereinigt. In alternativen Reinigungsprotokollen werden die Pellets in einem Puffer aufgenommen (Puffer A: 50 mM Tris-HCl pH 7.9, 50 mM Dextrose, 1 mM EDTA). Nach Zentrifugation werden die Zellen erneut aufgenommen, jedoch enthält Puffer A nun zusätzlich Lysozym (4 mg/ml). Nach Inkubation (15 min.) wird ein gleiches Volumen Puffer B zugegeben (B: 10 mM Tris-HCl (pH 7.9), 50 mM KCl, 1 mM EDTA, 1 mM PMSF, 0.5 % Tween 20, 0.5 % Nonidet P40) und die Lyse zur Inkubation bei 75°C für eine Stunde gegeben. Nach Zentrifugation wird der Überstand entnommen und die überexprimierten Proteine werden mittels (NH₄)₂SO₄ ausgefällt. Die Pellets werden nach Zentrifugation gesammelt und die Proteine mit Puffer A resuspendiert. Die resuspendierten Proteine werden gegen Aufbewahrungspuffer (50 mM Tris-HCl (pH 7.9), 50 mM KCl, 0.1 mM EDTA, 1 mM DTT, 0.5 mM PMSF, 50 % Glycerol) dialysiert und anschließend bei +4°C bis -70°C gelagert.

Um die Aktivität der Proteine zu testen, werden Reaktionen wie folgt zusammengesetzt. Aliquots der Proteine werden in unterschiedlichen Konfigurationen und Molaritäten vereint, Gleitklammerlader 1/2 mit Gleitklammer, koppelnder Untereinheit und Elongationsprotein 1, oder 5 Gleitklammerlader 1/2 mit Gleitklammer, mit und ohne koppelnder Untereinheit und Elongationsprotein 2, oder Gleitklammer, und Elongationsprotein 1 oder 2 und schließlich nur Elongationsprotein 1 oder 2; als Puffer dient der obige Aufbewahrungspuffer. DNA-Polymerisationsaktivität wird mittels Inkorporation von (methyl-³H) TTP in 10 Trichlorsäure-unlösliches Material oder mittels Inkorporation von Digoxigenin-dUTP in unmarkierte DNA-Doppelstrangbereiche mit freien internen 3'-Enden gemessen (Ishino, Y., Iwasaki, H., Fukui, H., Mineno, J., Kato, I., & Schinigawa, H. (1992) *Biochimie* 74, 131-136). Um die Prozessivität zu bestimmen, werden die obigen Proteingemische in Primer-15 Elongationsexperimenten verwendet. Ein M13-Einzelstrangtemplate wird in 10 mM Tris-HCl (pH 9.4) eingebracht und gemeinsam mit einem Universalprimer (New England Biolabs) (5'-FITC markiert) erhitzt (92°C) und abgekühlt (Raumtemperatur). Verdünnungsserien des so generierten Template-Primer-Gemisches werden in einer Reaktion bestehend aus Nukleotiden (etwa 200 20 µM bis 1 mM), Reaktionspuffer (Endkonzentration: 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 1.5-5 mM MgCl₂, ATP (0 mM – 200 mM) und proteinstabilisierenden Agenzien zusammengeführt und für 10 Minuten bei 37°C, 52°, 62°, 68°, 74°, und 78° inkubiert. Ein Aliquot wird zur Analyse auf einen Sequenzautomaten geladen (z.B. Alf, Pharmacie Biotech). Alternativ 25 kann eine qualitative Analyse der Steigerung der Prozessivität auch nach Maga G., Jónsson Z.O., Stucki M., Spadari S. und Hübscher U. (J. Mol. Biol. 1999; 285: 259-267, Dual Mode of Interaction of DNA-Polymerase ϵ with Proliferating Cell Nuclear Antigen in Primer Binding and DNA Synthesis) über den Nachweis einer Stimulation des Einbaus von Nukleotiden in 30 Doppelstrangbereiche mit freien internen 3'-Enden unter geeigneten Reaktionsbedingungen erfolgen (s. Fig. 23) Die obigen Proteingemische

dienen auch dazu Fidelity und Exonukleaseaktivität zu messen, wobei die in Kohler et al. (Proc. Natl. Acad. Sci. USA 88:7958-7962 (1991) oder Chase et al. (J. Biol. Chem., 249:4545-4552 (1972) beschriebenen Verfahren zur Anwendung kommen. Ebenso werden die Proteingemische in der PCR eingesetzt (Fig. 26; Methods in Molecular Biology Vol. 15 Humana Press Totowa, New Jersey, 1993, edited by Bruce A. White).

Beispiel 2: Trimerisierung von PCNA

In folgenden Experimenten wird gezeigt, daß rekombinantes *Archaeoglobus fulgidus* PCNA-Protein (AF 0335) unter nativen Bedingungen als Trimer vorliegt und somit eine für die Klammerbildung geeignete Struktur einnehmen kann.

Für das in Fig. 20 A abgebildete Experiment wurden 350 µl Ni-NTA Agarose Eluat von PCNA (s. Fig. 24) und 150 µl einer rohen DNA-Polymerase Fraktion (s. Fig. 25) mit Lagerpuffer ohne Glycerin (s. Fig. 25) auf 1 ml aufgefüllt und nach Vorschrift des Herstellers auf einer Superdex 200 HR (Pharmacia) FPLC Säule im gleichen Puffer die Proteine nach Molekulargewicht aufgetrennt. Es wurden während des gesamten Laufes Fraktionen zu je 1 ml gesammelt. Für das in Fig. 20 B. dargestellte Experiment wurden gleiche Mengen der gleichen Proteinfractionen in Storagepuffer ohne Glycerin mit 8 M Harnstoff auf 1 ml aufgefüllt, 10 Minuten bei 95°C denaturiert und anschließend die Proteine im gleichen Puffer wie für Fig. 20 A dargestellt nach Molekulargewicht aufgetrennt. Je 8 µl einzelner Fraktionen wurden anschließend auf einem NuPage Bis-Tris Gel aufgetrennt (NOVEX; s. Fig. 25) und nach Angaben des Herstellers mittels eines Blot-Modules (NOVEX) auf Nitrocellulosemembran geblottet. Der Nachweis von Proteinen mit His-Tag erfolgte nach Angaben der Hersteller mittels des RGS His Antibody (QIAGEN) und des DIG Luminiscent Detection Kit for Nucleic Acids (Boehringer Mannheim). Fig. 20 A zeigt Proteine mit His-Tag in den Fraktionen 15 (Spur 1), 17 (Spur 2), 19 (Spur 3), 21 (Spur 4), 23 (Spur 5), 25 (Spur 6) der ohne Harnstoff durchgeführten

Chromatographie. Fig. 20B zeigt Proteine mit His-Tag in den Fraktionen 10 (Spur 1), 11 (Spur 2), 12 (Spur 3), 13 (Spur 4), 14 (Spur 5), 15 (Spur 6), 16 (Spur 7), 17 (Spur 8) der in Anwesenheit Harnstoff durchgeführten denaturierenden Chromatographie. *Archaeoglobus fulgidus* DNA-Polymerase (AF0497) hat ein errechnetes Molekulargewicht von $M_r = 90$ kDa; *Archaeoglobus fulgidus*-PCNA (AF0335) hat ein errechnetes Molekulargewicht von $M_r = 27$ kDa. Wenn *Archaeoglobus fulgidus*-PCNA (AF 335), wie das homologe Protein aus Eukaryonten, als Homotrimer vorliegt, hat der native Faktor daher ein theoretisches Molekulargewicht von $M_r = 81$ kDa. Die in Fig. 20A gezeigten Ergebnisse bestätigen diese Annahme, nach der natives PCNA ein ähnliches Molekulargewicht aufweist, wie die DNA-Polymerase, für das rekombinante Protein: Beide Proteine eluieren unter nativen Bedingungen in den gleichen Fraktionen von der Gelfiltrationssäule (Fig 20 A, Spuren 1-3). Das Maximum der PCNA Menge eluiert dabei etwas später als das Maximum der DNA-Polymerasemenge und korreliert mit der etwas geringeren Größe des postulierten Trimers (81 kDa gegenüber 90 kDa). Die in Fig. 20 B gezeigten Daten belegen, daß diese Beobachtung auf einer Oligomerisierung des PCNA beruht, da unter denaturierenden Bedingungen, die Protein/Protein Interaktionen nicht zulassen, PCNA deutlich später von der Säule eluiert als die DNA-Polymerase (Fig. 20 B , Spuren 1-7), was dem geringeren Molekulargewicht des Monomers entspricht.

Beispiel 3: Isolierung, Bereitsstellung und Verwendung eines Elongationsproteins (*Pyrococcus horikoshii* DNA-Polymerase (PH1947)) zur Ausbildung des erfindungsgemäßen *in vitro*-Komplexes.

Das Elongationsprotein aus *Pyrococcus horikoshii* (*Pyrococcus horikoshii*-DNA-Polymerase (PH1947)) wurde in PCR-Reaktionen verwendet und führt zu effizienter Amplifikation eines spezifischen DNA Produktes.

1 (Spur 4), 2,5 (Spur 5) und 5 μ l (Spur 6) *Pyrococcus horikoshii* DNA-Polymerase (PH1947; Rohextrakt s. Fig. 25) wurden zum Aktivitätsvergleich mit 1 Einheit Taq Polymerase (Spur 2) und 1 μ l *Archaeoglobus fulgidus*-DNA-Polymerase (Af 497) Rohextrakt (s. Fig. 25) (Spur 3) individuell in Standard PCR-Reaktionen eingesetzt.; Spur 1 zeigt einen DNA-Größenmarker (New England Biolabs; Mischung aus 1 kb DNA Molekulargewichtsgroßenmarker und 100 bp DNA-Molekulargewichtsgroßenmarker). Jede Reaktion enthielt neben dem Enzym 5 ng template DNA (421 bp Rsa I-Fragment mit 10 Adaptoren in PCR 2.1 kloniert; (Kaiser C., v. Stein O., Laux G., Hoffmann M., Electrophoresis 1999; 20: 261-268, Functional Genomics in Cancer Research: Identification of Target Genes of the Epstein-Barr Virus Nuclear Antigen 2 by Subtractive cDNA Cloning and High-throughput Differential Screening using High-density Agarose Gels); je 0,2 mM dATP, dTTP, dCTP und dGTP; je 1,5 μ M der spezifischen Adaptor-Primer (Kaiser et al., (1999)); 50 mM KCl; 2 mM $MgCl_2$; 10 mM Tris HCl (pH 8,3 für Taq-Reaktionen; pH 7,5 für *Archaeoglobus fulgidus* und *Pyrococcus horikoshii*-Polymerase-Reaktionen) bei einem Gesamtvolumen von 50 μ l pro Reaktion. Die Proben durchliefen 40 Zyklen bestehend aus 30 s bei 95°C; 30 s bei 55°C und 120 s 20 bei 68°C. Anschließend wurden 5 μ l der Ansätze entnommen und auf einem 1% Agarosegel in 1 x TAE Puffer (40 mM Tris-Acetat; 20 mM Natriumacetat; 10 mM EDTA ; pH 7,2) mit 10V/cm aufgetrennt.

Beispiel 4: Verwendung eines Elongationsproteins

25

Die *Archaeoglobus fulgidus*-DNA-Polymerase AF 0497 kann PCR-Produkte ebenso effizient wie Taq-Polymerase generieren. 1 Einheit Taq-Polymerase und 1 μ l *Archaeoglobus fulgidus*-DNA-Polymerase AF 0497 Rohextrakt (s. Fig. 25) wurden zum Aktivitätsvergleich individuell in Standard-PCR- 30 Reaktionen eingesetzt. Jede Reaktion enthielt neben dem Enzym 20 ng M13 MP18 ssDNA; je 0,2 mM dATP, dTTP, dCTP und dGTP; je 1,5 μ M DNA

Primer	folgender	Nukleotidsequenz:
GGATTGACCGTAATGGGATAGGTTACGTT	(SEQ ID NO.: 47) bzw	
AGCGGATAACAATTTACACAGGAAACAG	(SEQ ID NO.: 48) in 50 mM	
KCl; 2 mM MgCl ₂ ; 10 mM Tris HCl (pH 8,3 für Taq Reaktionen; pH 7,5 für		
5 Archaeoglobus fulgidus Polymerase-Reaktionen) bei einem Gesamtvolumen		
von 50µl pro Reaktion. Die Proben durchliefen eine unterschiedliche Anzahl		
von Zyklen bestehend aus 30 s bei 95°C; 30 s bei 59°C und 60 s bei 68°C.		
Nach unterschiedlichen Zyklenzahlen (Z) wurden 5 µl der Ansätze		
entnommen und auf einem 1% Agarosegel in 1 x TAE Puffer (40 mM Tris-		
10 Acetat; 20 mM Natriumacetat; 10 mM EDTA ; pH 7,2) mit 10V/cm aufgetrennt		
(Siehe Fig. 22).		

Beispiel 5: Bereitsstellung und Verwendung eines erfindungsgemäßen thermostabilen *in vitro*-Komplexes:

15 In einem Endvolumen von 50 μ l waren unter anderem folgende Komponenten eines erfindungsgemäßen *in vitro*-Komplexes enthalten: 10 mM Tris/HCl pH 7,5; 50 mM KCl; 2 mM $MgCl_2$; 10 μ g BSA (kann auch weggelassen werden); 0,5 mM Digoxigenin-dUTP (DIG-dUTP, Boehringer Mannheim); 40 nM ; 0,5 μ g Poly dA/40 nM OligodT (20mer) Hybrid. Probe 1 ohne Elongationsprotein. Proben 2-12 enthielten außerdem je 3 μ l einer 1:1000 Verdünnung einer Fraktion von rekombinanter *Archaeoglobus fulgidus*-DNA-Polymerase (Elongationsprotein). Proben 3-7 sowie 8-12 außerdem 0,5; 1; 2; 4 und 8 μ l einer Fraktion von rekombinantem
25 Gleitklammerprotein aus *Archaeoglobus fulgidus*.

Die Proben wurden 30 Minuten bei 68°C inkubiert und Nukleinsäuren anschließend durch Präzipitation mit 3 Volumenteilen Ethanol / 3 M Natriumacetat pH 5,2 (30/1). Gefällt. Das Präzipitat wurde in 20 µl 100 mM Tris-HCl (pH 7,9) resuspendiert und je 10µl in einzelne Kavitäten einer 96-Well Silent Screen plate with Nylon 66 Biodyne B 0,45µm pore (Nalge Nunc)

aufgetropft und die Nukleinsäuren 10 Minuten bei 70°C auf der Membran fixiert. Die Detection von inkorporiertem Digoxigenin-dUTP (Boehringer Mannheim) erfolgte mittels des DIG Luminiscent Detection Kit for Nucleic Acids (Boehringer Mannheim). Zum Nachweis der Chemilumineszenz wurde
5 auf die Membran abschließend für 30 s ein Röntgenfilm aufgelegt. PCNA stimuliert die Einbindung von DIG-dUTP durch die DNA-Polymerase deutlich (vgl. Spuren 3-7 mit Spur 2). Die verwendete PCNA- Fraktion weist keine endogene Polymeraseaktivität auf (Spuren 8-12).

10 Beispiel 6: Amplifikation; Klonierung, Expression und Reinigung eines Gleitklammerproteins aus *Archaeoglobus fulgidus*:

Nach Amplifikation des erfindungsgemäßen Gleitklammerprotein Gens aus *Archaeoglobus fulgidus* wurde das Gen in die Expressionsvektoren pTrc99 und pQE30 kloniert. Die Expression, Reinigung und gelelektrophoretische
15 Auftrennung von *Archaeoglobus fulgidus*-Gleitklammerprotein (PCNA (AF 0335)) erfolgte, wie für das Elongationsprotein (die DNA-Polymerase AF 0497) in Fig. 25 dargestellt.

Beispiel 7: Bereitstellung eines Elongationsproteins

20 Expression und Reinigung der *Archaeoglobus fulgidus*-DNA-Polymerase:
pQE 30 Plasmid-DNA (QIAGEN) mit in korrektem Leseraster inseriertem Gen für das Elongationsprotein, die *Archaeoglobus fulgidus* DNA-Polymerase AF 0497, wurde nach Anleitung aus "The QIAexpressionist (dritte Auflage; QIAGEN)" in kompetente E.coli M15 [prep 4] transformiert.
25 Übertragung auf 1 L Kulturmedium und induzierte Proteinexpression erfolgte ebenfalls nach den dort vorgegebenen Schemata. Nach 16 Stunden Induktionszeit wurden die Bakterien 10 Minuten bei 5000 g sedimentiert. Zum Erhalt hochreiner Fraktionen wurde gemäß The QIAexpressionist (dritte Auflage; QIAGEN; protocol 8, protocol 11) verfahren. Lediglich die Elution
30 der gebundenen Proteine erfolgte mit 2 x 2 ml Elutionspuffer und nicht wie

dort beschrieben mit 4 x 0,5 ml Elutionspuffer. Zum Erhalt von Rohextrakten rekombinanter Proteine wurden die Bakteriensedimente alternativ mit 100 ml Puffer A (50 mM Tris-HCl pH 7,9; 50 mM Glukose; 1 mM EDTA) gewaschen und nach erneuter Zentrifugation in 50 ml Puffer A mit 4mg/ml Lysozym 5 resuspendiert. Nach 15 Minuten bei Raumtemperatur erfolgte die Zugabe von 50 ml Lysepuffer (10 mM Tris-HCl pH 7.9; 50 mM KCl; 1 mM EDTA; 0,5% Tween 20; 0,5% IGPAL) und E. coli-Proteine wurden durch 60-minütige Inkubation bei 75°C im Wasserbad denaturiert. Nach anschließender Zentrifugation für 15 Minuten bei 27000 g wurde der Überstand durch 10 langsame unter permanentem Rühren erfolgte Zugabe von 40 mg kristallinem Ammoniumsulfat pro ml Extrakt präzipitiert. Präzipitierte Proteine wurden 10 Minuten bei 27000 g sedimentiert und das Sediment in 20 ml Puffer A resuspendiert. Sowohl diese Rohextrakte als auch die Elutionsfraktionen aus der Ni NTA-Agarose wurden je 2-mal 8 Stunden 15 gegen mindestens je 50 Volumenteile Lagerpuffer (Puffer (10 mM Tris-HCl pH 7.9; 50 mM KCl; 1 mM EDTA; 50% Glycerin; 1 mM Dithiothreitol; 0,5 mM Phenylmethylsulfonylfluorid) dialysiert und bei -20°C aufbewahrt. Zur Analyse der Proteinzusammensetzung der erhaltenen Fraktionen wurden Teile davon nach Anweisungen des Herstellers auf NuPage™ 10% Bis-Tris 20 Gelen (NOVEX) elektrophoretisch aufgetrennt und die enthaltenen Proteine mit Coomassie brilliant blue angefärbt. Spur 1 zeigt einen Molekulargewichtsstandard ((BIO RAD Cat Nr 161-0317). Für Spur 2 wurden 500µl Bakterien direkt vor der Induktion sedimentiert, gemäß der Anweisung zum Lauf von NuPage Gelen behandelt und aufgetragen. Spur 3 zeigt die 25 gleiche Menge Bakterien 16 Stunden nach der Elution. Spuren 4 und 5 zeigen je 8µl der beiden Eluate der Ni-Nta-Agarosesäule nach der Dialyse. Spur 6 zeigt 8 ml eines dialysierten Grobextraktes. Durch Reinigung über Ni-NTA-Agarose werden hochreine Fraktionen der *Archaeoglobus fulgidus* DNA-Polymerase erhalten (Spuren 4 und 5). Aber bereits im Grobextrakt 30 stellt dieses Enzym das dominierende Polypeptid dar.

Beispiel 8: Verwendung eines erfindungsgemäßen *in vitro* Komplexes in der PCR

Der Einsatz von *Archeoglobus fulgidus* PCNA (AF 0335) in PCR Reaktionen mit *Archeoglobus fulgidus* DNA-Polymerase (AF 0497) führt zu einer effizienteren Amplifikation eines spezifischen DNA-Produkts im Vergleich zu einer PCR-Reaktion ohne PCNA. Die Auswertung der amplifizierten DNA-Mengen laut AIDA (Advanced Image Data Analyzer, Software Version AIDA 2.1, Firma Raytest) zeigt einen Hintergrund von 94, einen Wert von 104,9 für Spur 1 und einen Wert von 228,4 für Spur 2. Zieht man von den Werten aus den Spuren 1 und 2, den Hintergrund ab, folgt daraus eine 12,3-fache prozessive Stimulation durch PCNA in der Reaktion dessen Ergebnis in Spur 2 gezeigt ist.

15 0,3 µl *Archeoglobus fulgidus*-DNA-Polymerase (7.5 µg/µl Proteinkonzentration von AF497; Rohextrakt siehe Fig. 25) wurden zur Analyse einer Stimulation der PCR-Aktivität durch *Archeoglobus fulgidus*-PCNA (NiNTA Eluat; Fig.24) mit 0 µl und 0,01 µl individuell in PCR Reaktionen eingesetzt. Jede Reaktion enthielt neben dem Enzym und 0,05 µl
20 (2.8 µg) des Gleitklammerproteins von *Archeoglobus fulgidus* (homolog zu proliferating-cell-nuclear antigen, PCNA) eines mittels PCR amplifizierten und nicht aufgereinigten PCNA-Gen-Fragments (PCR Reaktion zur spezifischen Amplifikation des PCNA-Fragments: 0,5 µl *Archeoglobus fulgidus* Polymerase; je 0,2 mM dATP, dTTP, dCTP und dGTP; je 1,5 µM der
25 spezifischen Primer (5'-ACG CGC GGA TCC ATA GAC GTC ATA ATG ACC GG-3' (SEQ ID No.: 49); 5'-TAC GGG GTA CCC GAG CCA AAA TTG GGT AAA G-3' (SEQ ID No.:50); 50 mM KCl; 2 mM MgCl₂; 10 mM Tris-HCl (pH 7,5) sowie 50 ng eines pQE30 Plasmids, das die kodierenden Sequenzen von *Archeoglobus fulgidus* PCNA in die BamHI- und Kpn I-
30 Restriktionsstellen inseriert trägt, bei einem Gesamtvolumen von 50 µl pro

Reaktion und bei einer Zyklenzahl von 40 bestehend aus 30 s bei 95°C; 30 s bei 61°C; 240 s bei 68°C); je 0,2 mM dATP, dTTP, dCTP und dGTP; je 1,5 µM der spezifischen Primer

(5'-ACG CGC GGA TCC ATA GAC GTC ATA ATG ACC GG-3' (SEQ ID No.: 51); und 5'-TAC GGG GTA CCC GAG CCA AAA TTG GGT AAA G-3' (SEQ ID No.: 52); 50 mM KCl; 2 mM MgCl₂; 10 mM Tris HCl pH 7,5 bei einem Gesamtvolumen von 50 µl pro Reaktion. Die Proben durchliefen 40 Zyklen bestehend aus 30s bei 95°C; 120s bei 61°C; 240s bei 68°C. Anschließend wurden 5 µl der Ansätze entnommen und auf ein 1 % Agarosegel in 1 x TAE Puffer (40 mM Tris-Acetat; 20 mM Natriumacetat; 10 mM EDTA; pH 7,2) mit 10 V/cm aufgetrennt.

Beispiel 9: Y2H-Experimente:

Interaktionen von Proteinen des erfindungsgemäßen Komplexes aus *Archaeoglobus fulgidus* wurden mit dem Y2H-System gezeigt. Die kodierenden Bereiche von Genen aus *Archaeoglobus fulgidus*, deren Genprodukte im erfindungsgemäßen thermostabilen *in vitro*-Komplex verwendeten werden, wurden mittels PCR amplifiziert, in die Vektoren pGBT9 (vertikale Spalten der Fig. 27) und pGAD424 (horizontale Reihen der Fig. 27) kloniert und durch gap-repair in Hefe PJ69-4a (für pGAD424) und PJ69-4alpha (für pGBT9) als Hybridproteine zur Expression gebracht. Ebenso wurde eine positive Kontrolle mittels PCR amplifiziert, in die Vektoren pGBT9 (siehe auch vertikale Spalten der Fig. 27) und pGAD424 (siehe auch horizontale Reihen der Fig. 27) kloniert und durch gap-repair den Hefestamm PJ69-4a (für pGAD424) und PJ69-4alpha (für pGBT9) als Hybridproteine zur Expression gebracht. Diploide Zellen, die beide Vektoren enthalten, wurden durch Paarung entsprechend dem gezeigten Raster in Fig. 27 erzeugt. Die Expression von drei unabhängigen Reportern (*HIS3*, *ADE2* und *MEL1*) wurde gemessen. Gezeigt ist in Fig. 27 das Wachstum auf Medium ohne Histidin und Adenin. In diesem Experiment wachsen jene Zellen, die beide Vektoren tragen und in dessen zudem die

Expressionsprodukte dieser beiden Vektoren einander binden. In Folge der Bindung der Expressionsprodukte wird Transkription der Reportergene initiiert, so daß die Histidin- und Adenin-Auxotrophie aufgehoben wird und diese Zellen in der Lage sind in besagtem Medium zu wachsen.

5

Alle hier positiven Klone waren auch positiv bezüglich der Expression des *MEL1*-Gens. Die Handhabung des yeast two-hybrid (Y2H)systems erfolgte gemäß der Anleitung der Firma Clontech (yeast protocols handbook, PT3024-1). Fig. 27 zeigt eine Bindung zwischen den Proteinen der
10 Positivkontrolle, dem Elongationsprotein und dem Gleitklammerprotein, dem Gleitklammerprotein und dem Gleitklammerprotein und dem Kopplungsprotein und dem Gleitklammerprotein.

Beispiel 10: Verwendung eines erfindungsgemäßen thermostabilen in

15 *vitro*-Komplexes:

Der erfindungsgemäße thermostabile *in vitro*-Komplex läßt sich sehr gut bei der Amplifikation von genomischen DNA Fragmenten einsetzen. Hierbei kommt es auch bei Verwendung von geringen Mengen eines Templates oder
20 eines Elongationsproteins zu effizienter Amplifikation. In Beispiel 10 wurde ein Abschnitt des humanen Kollagen-Gens amplifiziert, einmal unter Verwendung des erfindungsgemäßen thermostabilen *in vitro*-Komplexes und einmal unter Verwendung eines Elongationsproteins alleine. In einem Gesamtvolumen von 50 µl wurden 0,5 µl Nukleotidmix (25 mM
25 Ausgangslösung beinhalten jedes Nukleotid A, C, G und T), 0,2 µl jedes Primers (100 pmol/µl Ausgangslösung "Collagen Forward" 5'-TAA AGG GTC ACC GTG GCT TC-3' (SEQ ID NO.: 53), 100 pmol/µl Ausgangslösung "Collagen Reverse" 5'- CGA ACC ACA TTG GCA TCA TC-3' SEQ ID NO.: 54), 0,8 µl DNA (200 ng/µl human genomic DNA von Boehringer Mannheim),
30 5 µl 10 x PCR Puffer (pH 7.5) (100mM Tris-HCl pH7.5, 500 mM KCl, 15 mM

MgCl₂), 1 µl Elongatinsprotein (AF1722, 7.5 µg/µl Proteinkonzentration) und 8 µl Gleitklammerprotein (AF 0335, Proteinkonzentration 0,3 µg/µl) eingesetzt und in einer PCR-Maschine dem folgenden Zyklus unterworfen, anfangs 5 min. bei 95°C und dann 30-mal, 30 s bei 95°C, 30 s bei 59°C und 5 abschließend 6 min. bei 68°C. Die Fig. 28 zeigt deutlich den Vorteil der Verwendung des erfindungsgemäßen thermostabilen *in vitro*-Komplexes.

Die in der vorstehenden Beschreibung sowie in den Ansprüchen offenbarten Merkmale der Erfindung können sowohl einzeln als auch in beliebiger Kombination für die Verwirklichung der Erfindung in ihren verschiedenen Ausführungsformen wesentlich sein.

Patentansprüche

1. Thermostabiler *in vitro*-Komplex zur Template-abhängigen Elongation
5 von Nukleinsäuren, umfassend ein thermostabiles
 Gleitklammerprotein und ein thermostabiles Elongationsprotein.
2. Thermostabiler *in vitro*-Komplex nach Anspruch 1, dadurch
 gekennzeichnet, daß das Gleitklammerprotein mit einem
10 Elongationsprotein verbunden ist.
3. Thermostabiler *in vitro*-Komplex nach Anspruch 1, dadurch
 gekennzeichnet, daß ein Gleitkammerprotein mit einem
15 Elongationsprotein direkt verbunden ist.
4. Thermostabiler *in vitro*-Komplex nach Anspruch 1 oder 2,
 dadurch gekennzeichnet,
 dass das Gleitklammerprotein und das Elongationsprotein über ein
20 Kopplungsprotein verbunden sind.
5. Thermostabiler *in vitro*-Komplex nach einem der Ansprüche 1 bis 4,
 dadurch gekennzeichnet,
 dass das Gleitklammerprotein und/oder das Elongationsprotein aus
25 Archaeobakterien stammen.
6. Thermostabiler *in vitro*-Komplex nach einem der Ansprüche 1 bis 5,
 dadurch gekennzeichnet,

dass das Gleitklammerprotein eine ringförmige Struktur aufweist, welche die Template-Nukleinsäurestränge ganz oder teilweise umschließt.

- 5 7. Thermostabiler *in vitro*-Komplex nach einem der Ansprüche 1 bis 6, dadurch gekennzeichnet, dass das Gleitklammerprotein eine oder beide der folgenden Aminosäurekonsensussequenzen umfaßt:

SEQ ID No.:39

- 10 [GAVLIMPFW]-D-X-X-X-[GAVLIMPFW]-X-X-[GAVLIMPFW]-X-
[GAVLIMPFW]-X-[GAVLIMPFW]-X-X-X-X-F-X-X-Y-X-X-D
und / oder

SEQ ID No.: 40

[GAVLIMPFW]-X(3)-L-A-P-[KRHDE]-[GAVLIMPFW]-E.

15

8. Thermostabiler *in vitro*-Komplex nach einem der Ansprüche 1 bis 7, dadurch gekennzeichnet,

- dass das Gleitklammerprotein zu der menschlichen (Eukaryonten) PCNA Aminosäuresequenz (SEQ ID NO. 11) auf einer Länge von
20 mindestens 100 Aminosäuren bei einem Sequenzalignment eine mindestens 20 %ige Sequenzidentität aufweist und/oder das Gleitklammerprotein zu der bakterielle β -clamp Sequenz aus E.coli (Eubakteria) (SEQ ID NO. 35) auf einer Länge von mindestens 100 Aminosäuren bei einem Sequenzalignment eine mindestens 20 %ige
25 Sequenzidentität aufweist und/oder das Gleitklammerprotein zu der Aminosäuresequenz des PCNA-Homologen aus *Archaeoglobus fulgidus* (SEQ ID NO. 12) auf einer Länge von mindestens 100 Aminosäuren bei einem Sequenzalignment eine mindestens 20 %ige
30 Sequenzidentität aufweist.

9. Thermostabiler Komplex nach einem der vorhergehenden Ansprüche,
dadurch gekennzeichnet,
dass das Gleitklammerprotein mit einem aus einem Alignment aus Fig.
12 generierten Hidden Markow Modell einen Score von mindestens 20
5 ergibt und/oder wobei das Gleitklammerprotein mit einem aus einem
Alignment aus Fig. 13 generierten Hidden Markow Modell einen Score
von mindestens 25 ergibt.
10. Thermostabiler *in vitro*-Komplex nach einem der vorhergehenden
10 Ansprüche, dadurch gekennzeichnet, daß das Gleitklammerprotein ein
Gleitklammerprotein ist, das aus einem Organismus stammt, der
ausgewählt ist aus der Gruppe, die *Archaeoglobus fulgidus*,
Methanoccus jannaschii, *Pyrococcus horikoshii*, *Methanobacterium*
thermoautotrophicus, *Aquifex aeolicus* und *Carboxydothemus*
15 *hydrogenofhormans* umfaßt.
11. Thermostabiler *in vitro*-Komplex nach einem der vorhergehenden
Ansprüche,
dadurch gekennzeichnet,
20 dass das Gleitklammerprotein ausgewählt ist aus der Gruppe, die
AF0335 aus *Archaeoglobus fulgidus*, MJ0247 aus *Methanococcus*
jannaschii, PHLA008 aus *Pyrococcus horikoshii*, MTH1312 aus
Methanobacterium thermoautotrophicus sowie AE000761_7 aus
Aquifex aeolicus umfaßt.
25
12. Thermostabiler *in vitro*-Komplex nach einem der vorhergehenden
Ansprüche,
dadurch gekennzeichnet,

dass das Elongationsprotein eine 5'-3'-Polymeraseaktivität und/oder eine Reverse-Transkriptaseaktivität aufweist.

13. Thermostabiler *in vitro*-Komplex nach Anspruch 12,
5 **dadurch gekennzeichnet,**

dass das Elongationsprotein mindestens eine der folgenden Konsensussequenzen umfaßt und an nicht mehr als vier Positionen von dieser Sequenz abweicht:

SEQ ID NO. 44:

- 10 D-[GAVLIMPFW]-[GAVLIMPFW]-X-X-Y-N-X-X-X-F-D-X-P-Y-
[GAVKUNOFW]-X-X-R-A

SEQ ID NO. 45

A-[GAVLIMPFW]-R-T-A-[FAVLIMPFW]-A-[GAVLIMPFW]-[GAVLIMPFW]-T-
E-G-[GAVLIMPFW]-V-X-A-P-[GAVLIMPFW]-E-G-I-A-X-V-[KRHDE]-I

- 15 SEQ ID NO. 46

[GAVLIMPFW]-P-V-G-[GAVLIMPFW]-G-R-G-S-X-[GAVLIMPFW]-G-S-
[GAVKUNOFW]-V-A-X-A-[GAVLIMPFW]-X-I-T-D-[GAVKUNOFW]-D-P-
[GAVLIMPFW]-X-X-X-[GAVLIMPFW]-L-F-E-R-F-L-N-P-E-R-[GAVLIMPFW]-
S-M-P-D.

20

14. Thermostabiler *in vitro*-Komplex nach Anspruch 12,

dadurch gekennzeichnet,

dass das Elongationsprotein eine zu der menschlichen (Eukaryonten) Aminosäuresequenz (SEQ ID NO. 22) auf einer Länge von mindestens 200

- 25 Aminosäuren bei einem Sequenzalignment eine mindestens 20 %ige Sequenzidentität aufweist und/oder zur archaebakteriellen Aminosäuresequenz (SEQ ID NO. 27) auf einer Länge von mindestens 400 Aminosäuren bei einem Sequenzalignment eine mindestens 25 %ige Sequenzidentität aufweist und/oder zur eubakteriellen Aminosäuresequenz
30 (SEQ ID NO. 37) auf einer Länge von mindestens 300 Aminosäuren bei

einem Sequenzalignment eine mindestens 25 %ige Sequenzidentität aufweist.

15. Thermostabiler *in vitro*-Komplex nach Anspruch 12,
5 **dadurch gekennzeichnet,**
dass das Elongationsprotein mit einem aus einem Alignment aus Fig. 17 generierten Hidden Markow Modell einen Score von mindestens 20 ergibt und/oder das Elongationsprotein mit einem aus einem Alignment aus Fig. 18 generierten Hidden Markow Modell einen Score von mindestens 35 ergibt und/oder das Elongationsprotein mit einem aus einem Alignment aus Abbildung 19 generierten Hidden Markow Modell einen Score von mindestens 20 ergibt.
16. Thermostabiler *in vitro*-Komplex nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet,** daß das Elongationsprotein ein Elongationsprotein ist, das aus einem Organismus stammt, der aus der Gruppe ausgewählt ist, die *Archaeoglobus fulgidus*, *Methanococcus jannaschii*, *Pyrococcus horikoshii*, *Methanobacterium thermoautotrophicus*, *Pyrococcus furiosus* und *Carboxydotherrnus hydrogenoformans* umfaßt.
17. Thermostabiler *in vitro*-Komplex nach einem der vorhergehenden Ansprüche,
dadurch gekennzeichnet,
25 dass das Elongationsprotein ausgewählt ist aus der Gruppe, die AF0497 oder AF1722 aus *Archaeoglobus fulgidus*, MJ0885 oder MJ1630 aus *Methanococcus jannaschii*, PHBT047 oder PHBN021 aus *Pyrococcus horikoshii*, MTH1208 oder MTH1536 aus *Methanobacterium thermoautotrophicus* sowie PFUORF3 aus
30 *Pyrococcus furiosus* umfaßt.

18. Thermostabiler *in vitro*-Komplex nach einem der vorhergehenden Ansprüche,
dadurch gekennzeichnet,
5 dass das Kopplungsprotein die folgende Konsensussequenz beinhaltet und an nicht mehr als vier Positionen von dieser Sequenz abweicht:
(SEQ ID NO.43:)
[FL]-[GAVLIMPFW]-X-X-[GAVLIMPFW]-X-G-X(13)-[GAVLIMPFW]-X-[YR]-
10 [GAVLIMPFW]-X-[GAVLIMPFW]-A-G-[DN]-[GAVLIMPFW]-[GAVLIMPFW]-[DS].
19. Thermostabiler *in vitro*-Komplex nach einem der vorhergehenden Ansprüche,
15 dadurch gekennzeichnet,
dass das Kopplungsprotein zu der menschlichen (Eukaryonten) Aminosäuresequenz (SEQ ID NO: 16) auf einer Länge von mindestens 150 Aminosäuren bei einem Sequenzalignment eine mindestens 18 %ige Sequenzidentität aufweist.
20
20. Thermostabiler *in vitro*-Komplex nach einem der vorhergehenden Ansprüche,
dadurch gekennzeichnet,
25 dass das Kopplungsprotein mit einem aus einem Alignment aus Abbildung 16 generierten Hidden Markow Modell einen Score von mindestens 10 ergibt.
21. Thermostabiler *in vitro*-Komplex nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß das Kopplungsprotein ein
30 Kopplungsprotein ist, das aus einem Organismus stammt, der aus der

Gruppe ausgewählt ist, die *Archaeoglobus fulgidus*, *Methanococcus jannaschii*, *Pyrococcus horikoshii*, *Methanobacterium thermoautotrophicus*, *Pyrococcus furiosus* und *Carboxydothemus hydrogenophormans* umfaßt.

5

22. Thermostabiler *in vitro*-Komplex nach einem der vorhergehenden Ansprüche,

dadurch gekennzeichnet,

- 10 dass das Kopplungsprotein ausgewählt ist aus der Gruppe, die AF1790 aus *Archaeoglobus fulgidus*, MJ0702 aus *Methanococcus jannaschii*, PHBN023 aus *Pyrococcus horikoshii*, MTH1405 aus *Methanobacterium thermoautotrophicum* sowie PFUORF2 aus *Pyrococcus* umfaßt.

- 15 23. Thermostabiler *in vitro*-Komplex nach einem der vorhergehenden Ansprüche,

dadurch gekennzeichnet,

dass der Komplex mit einem Protein assoziiert ist, welches als Gleitklammerlader fungiert.

20

24. Thermostabiler Komplex nach einem der vorhergehenden Ansprüche,

dadurch gekennzeichnet,

dass der Komplex mit ATP oder einem anderen Kofaktor assoziiert vorliegt.

25

25. Rekombinante DNA-Sequenz,

dadurch gekennzeichnet,

dass sie für einen thermostabilen *in vitro*-Komplex gemäß einem der Ansprüche 1 bis 24.

30

26. Vektor,
dadurch gekennzeichnet,
dass er eine für ein Gleitklammerprotein und ein Kopplungsprotein
und/oder ein Elongationsprotein codierende, rekombinante DNA-
Sequenz enthält.
27. Vektor nach Anspruch 26,
dadurch gekennzeichnet,
dass er zusätzlich mindestens eine weitere DNA-Sequenz mit
mindestens einer geeigneten Restriktionsschnittstelle zur Insertion
weiterer DNA-Sequenzen in einer solchen Anordnung enthält, dass
sich ein Fusionsprotein aus Gleitklammerprotein und dem
Expressionsprodukt der weiteren DNA-Sequenzen ergibt.
28. Vektor nach Anspruch 26 oder 27,
dadurch gekennzeichnet,
dass er geeignete, die Expression der DNA-Sequenz(en) steuernde
Promotor- und/oder Operatorbereiche enthält.
29. Vektor nach Anspruch 28,
dadurch gekennzeichnet,
dass er mehrere Promotor- und/oder Operatorbereiche zur
getrennten Expression mehrerer DNA-Sequenzen enthält.
30. Vektor nach einem der Ansprüche 26 bis 29,
dadurch gekennzeichnet,
dass er reprimier- und/oder induzierbare Promotor- und/oder
Operatorbereiche enthält.

31. Vektor nach einem der Ansprüche 26 bis 30,
dadurch gekennzeichnet,
dass er eine DNA-Sequenz nach Anspruch 25 enthält.
- 5 32. Wirtszelle,
dadurch gekennzeichnet,
dass sie mit einem oder mehreren Vektoren nach einem der
Ansprüche 26 bis 31 transformiert ist.
- 10 33. Verfahren zur Herstellung eines thermostabilen *in vitro*-Komplexes
gemäß einem der Ansprüche 1 bis 24,
dadurch gekennzeichnet,
dass eine entsprechende rekombinante DNA-Sequenz gemäß
Anspruch 25 oder einen oder mehrere der Vektoren gemäß einem der
15 Ansprüche 26 bis 31 in eine Wirtszelle einbringt, die Expression der
Proteine bewirkt und diese aus dem Kulturmedium oder nach
Zellaufschluss isoliert und gegebenenfalls noch mit weiteren
Komplexbestandteilen koppelt.
- 20 34. Verfahren zur Template-abhängigen Elongation von Nukleinsäuren,
wobei die Nukleinsäure nötigenfalls denaturiert wird, mit mindestens
einem Primer unter Hybridisierungsbedingungen versehen wird, wobei
der Primer hinreichend komplementär zu einem flankierenden Bereich
einer gewünschten Nukleinsäuresequenz des Template-Strangs ist und
25 mit Hilfe einer Polymerase in Anwesenheit von Nukleotiden eine Primer-
Elongation erfolgt,
dadurch gekennzeichnet,
dass als Polymerase ein thermostabiler *in vitro*-Komplex gemäß einem
der Ansprüche 1 bis 24 verwendet wird.

35. Verfahren nach Anspruch 34,
dadurch gekennzeichnet,
dass man zur Amplifikation von DNA-Sequenzen zwei die gewünschte
5 Nukleinsäuresequenz flankierende Primer und Deoxynukleotide
und/oder Derivate davon und/oder Ribonukleotide und/oder Derivate
davon verwendet.
36. Verfahren nach Anspruch 35,
10 **dadurch gekennzeichnet,**
dass man eine Polymerase-Ketten-Reaktion durchführt.
37. Verfahren nach Anspruch 34,
dadurch gekennzeichnet,
15 dass man zur Reversen-Transkription von RNA in DNA einen
thermostabilen *in vitro*-Komplex gemäß einem der Ansprüche 1 bis 24
einsetzt, dessen Elongationsprotein Reverse-Transkriptase-Aktivität
aufweist.
- 20 38. Verfahren nach einem der Ansprüche 34 bis 37
dadurch gekennzeichnet,
dass man zur Sequenzierung von Nukleinsäuren ausgehend von einem
Primer, der zu einem der zu sequenzierenden Nukleinsäure benachbarten
Bereich komplementär ist, eine Template-abhängige Elongation bzw.
25 Reverse-Transkription unter Verwendung von Deoxynukleotiden und
Dideoxynukleotiden oder deren jeweiligen Derivaten gemäß der Methode
von Sanger durchführt.

39. Verfahren nach einem der Ansprüche 34 bis 38,
dadurch gekennzeichnet,
dass man bei der Elongation der Nukleinsäuren Markierungen einfügt.
- 5 40. Verfahren nach Anspruch 39,
dadurch gekennzeichnet,
dass man markierte Primer und/oder markierte Deoxynukleotide
und/oder deren Derivate und/oder markierte Dideoxynukleotide
und/oder deren Derivate und/oder markierte Ribonukleotide und/oder
10 deren Derivate einsetzt.
41. Verfahren zur Markierung von Nukleinsäuren durch Erzeugung
einzeln Brüche in Phosphodiesterbindungen der Nukleinsäurekette
und Ersatz eines Nukleotids an den Bruchstellen durch ein markiertes
15 Nukleotid mit Hilfe einer Polymerase,
dadurch gekennzeichnet,
dass als Polymerase ein thermostabiler *in vitro*-Komplex nach einem
der Ansprüche 1 bis 24 eingesetzt wird.
- 20 42. Reagenzien-Kit zur Elongation und/oder Amplifikation und/oder
reversen Transkription und/oder Sequenzierung und/oder Markierung
von Nukleinsäuren, enthaltend in einem oder mehreren getrennten
Behältern
- 25 a) einen thermostabilen *in vitro*-Komplex nach einem der
Ansprüche 1 bis 24 oder
b) einen thermostabilen *in vitro*-Komplex nach einem der
Ansprüche 1 bis 24 und separat davon ein Polymeraseaktivität
aufweisendes Elongationsprotein,

sowie gegebenenfalls Primer, Puffersubstanzen, Nukleotide, ATP, einen oder mehrere andere Kofaktoren und/oder Pyrophosphatase.

43. Kit nach Anspruch 42,

5 **dadurch gekennzeichnet,**
dass er zur Amplifikation von Nukleinsäuren neben den Komponenten a) oder b), welche 5'-3'-Polymeraseaktivität aufweisen, Deoxynukleotide und/oder Derivate davon enthält.

10 44. Kit nach Anspruch 42,

dadurch gekennzeichnet,
dass er zur reversen Transkription die Komponenten a) oder b) enthält, welche reverse Transkriptase-Aktivität aufweisen, sowie Deoxynukleotide und/oder Derivate davon.

15

45. Kit nach einem der Ansprüche 42 bis 44,

dadurch gekennzeichnet,
dass er zur Sequenzierung neben Deoxynukleotiden oder Ribonukleotiden und/oder deren Derivaten, Dideoxynukleotide und/oder deren Derivate enthält.

20

46. Kit nach einem der Ansprüche 42 bis 45,

dadurch gekennzeichnet,
dass er Primer und/oder Deoxynukleotide und/oder Dideoxynukleotide in markierter Form enthält.

25

Fig. 1:

Komponenten des Elongations- Inkubationszyklus	Eukaryota Homo sapiens	Archaeon Fulgoidus (%Identität Alignmentslänge) ¹	Archaeon Methanococcus Jannaschii	Archaeon Pyrococcus horikoshii (%Identität) ²	Archaeon Methanobact. Thermoautot.	Archaeon Pyrococcus Furiosus (%Identität) ²
Gleitkammer- lader I	AC11_HUMAN [SEQ ID NO:1]	AF2060 (39% 307) [SEQ ID NO:2]	MJ1422 (32% 173) [SEQ ID NO:3]	PHB012 (39% 269) [SEQ ID NO:4]	MTH0241 (35% 317) [SEQ ID NO:5]	-----
	AC12_HUMAN [SEQ ID NO:32]	AF2060 (45%) [SEQ ID NO:2]				
	AC13_HUMAN [SEQ ID NO:33]	AF2060 (27%) [SEQ ID NO:2]				
	AC14_HUMAN [SEQ ID NO:34]	AF2060 (42%) [SEQ ID NO:2]				
Gleitkammer- lader II	AC15_HUMAN [SEQ ID NO:6]	AF1195 (29% 276) [SEQ ID NO:7]	MJ0884 (25% 571) [SEQ ID NO:8]	PHB013 (23% 429) [SEQ ID NO:9]	MTH0240 (26% 308) [SEQ ID NO:10]	-----
Gleitkammer	PCNA_HUMAN [SEQ ID NO:11]	AF0335 (24% 257) [SEQ ID NO:12]	MJ0247 (25% 256) [SEQ ID NO:13]	PHLA008 (25% 245) [SEQ ID NO:14]	MTH1312 (28% 260) [SEQ ID NO:15]	-----
koppelnde Untereinheit	DPD2_HUMAN [SEQ ID NO:16]	AF1790 (23% 211) [SEQ ID NO:17]	MJ0702 (19% 340) [SEQ ID NO:18]	PHB023 (21% 401) ³ [SEQ ID NO:19]	MTH1405 (21% 265) [SEQ ID NO:20]	PFUORF2 (43% 378) ² [SEQ ID NO:21]
Elongationsprot.I (Polymerase)	DPOD_HUMAN [SEQ ID NO:22]	AF0497 (25% 742) [SEQ ID NO:23]	MJ0885 (28% 277) [SEQ ID NO:24]	PHB047 (29% 439) [SEQ ID NO:25]	MTH1208 (26% 650) [SEQ ID NO:26]	-----
Elongationsprot.II (Pol.Aktivität)	-----	AF1722 [SEQ ID NO:27]	MJ1630 (50% 1160) ² [SEQ ID NO:28]	PHB021 (47% 974) ² [SEQ ID NO:29]	MTH1536 (49% 1139) ² [SEQ ID NO:30]	PFUORF3 (50% 1198) ² [SEQ ID NO:31]

Fig. 2:

Komponenten des Elongations- holoenzymes	Eubakteria Escherichia coli	Aquifex Aeolicus (%Identität Alignmentlänge) ¹
Gleitklammer	DP3B_ECOLI [SEQ ID NO:35]	AASEQ93 (28% 370) [SEQ ID NO:36]
Elongationsprotein	DP3A_ECOLI [SEQ ID NO:37]	AASEQ50 (37% 1154) [SEQ ID NO:38]

Fig. 3:

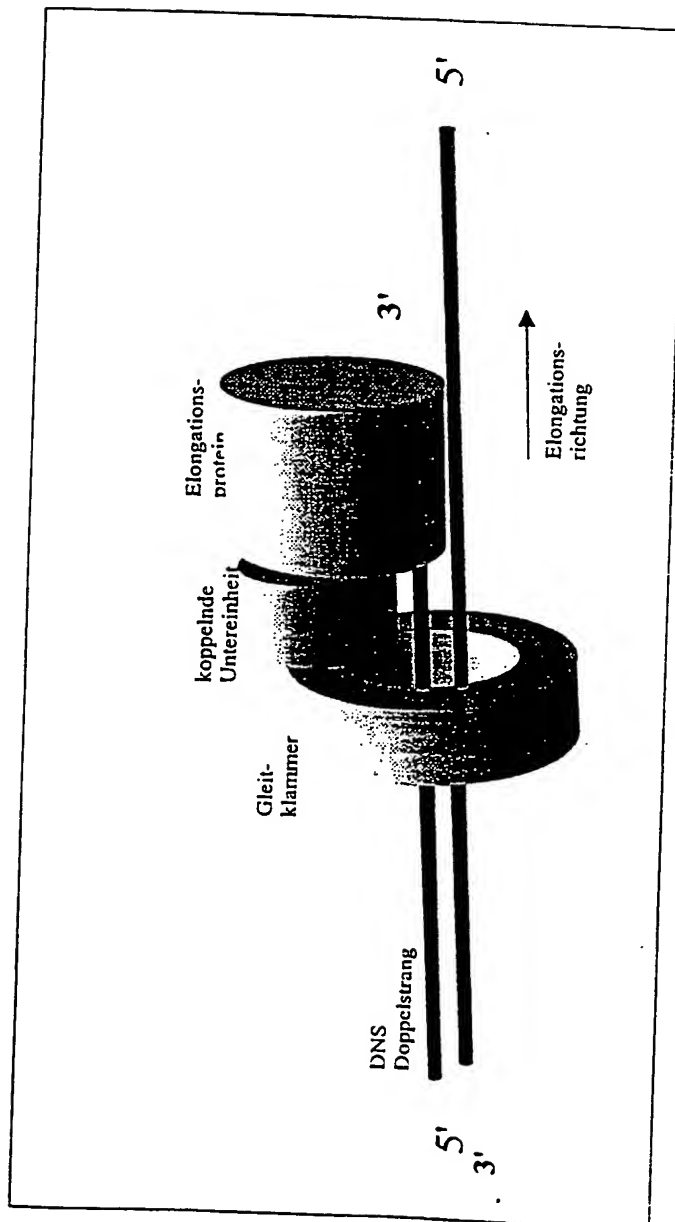


Fig. 4:

4/28

Gleitklammer Region 1:

PCNA_HUMAN : MDSSHVSLVQLTLRSEGFDTYRCD
PCNA_METJA : MDP SHVALVSLEIPRLAFEEYEAD
MTH1312 : LDRSHITYVHLELKAELFDEYVCD
PHLA008 : MDPSRVVLIDLNLPS SIFSKYEVD
AF0335 : VDPANVAMVIVDIPKDSFEVYNID

Konsensus : \$DXXX\$XX\$X\$X\$XXXXFXXYXXD

Pattern : [GAVLIMPFW]-D-X-X-X-[GAVLIMPFW]-X-X-[GAVLIMPFW]-X-
[GAVLIMPFW]-X-[GAVLIMPFW]-X-X-X-X-F-X-X-Y-X-X-D

Gleitklammer Region 2:

PCNA_HUMAN : LKYYLAPKIE
PCNA_METJA : LTFL LAPRIE
MTH1312 : LSFL LAPRIE
PHLA008 : LTFL LAPRVE
AF0335 : VEYILAPRIE

Konsensus : \$XXXLAP&\$E

Pattern : [GAVLIMPFW]-X(3)-L-A-P-[KRHDE]-[GAVLIMPFW]-E

5/28

Fig. 5:

Koppelnde Untereinheit:

PHBN023 : FLEWLNGYVESKEEEEIVSRIRYLIIAGDVVD
 PfuORF2 : FLEWLNGNVETKEEEEIVSRVKYLIAGDVVD
 DPD2_HUMAN: LVDVVTGQLGDEGEQCSAAHVSRVILAGNLLS
 MTH1405 : FVKWINGDFGSEEQSLAADVKYLVVAGDIVD
 AF1790 : FVRWLKGEVGGKKSQNLAEKVYIVIAGDIVD
 MJ0702 : FIRFLNGDVDNELEEKVVSRLKYICIAGDLVD

Konsensus : F\$XX\$XGXXXXXXXXXXXXX\$XY\$X\$AGD\$SD
 L R N S

Pattern : [FL] - [GAVLIMPFW] - X - X - [GAVLIMPFW] - X - G - X (13) -
 [GAVLIMPFW] - X - [YR] - [GAVLIMPFW] - X - [GAVLIMPFW] - A - G - [DN] -
 [GAVLIMPFW] - [GAVLIMPFW] - [DS]

Fig. 6:

Gleitklammerlader:

AC11_HUMAN : CNYLSKIIPALQSRCTRFRFGPL
 AF2060 : CNYVSRIIEPIQSRCAVFRFKPV
 MTH0241 : CNYSSKIIDPIQSRCAIFRFLPL
 PHBN012 : CNYSSKIIEPIQSRCAIFRFRPL
 MJ1422 : CNYPSKIIPPIQSRCAVFRFSPL

Konsensus : CNYXS&IIX\$QSRCXXFRFXP\$

Pattern : C - N - Y - X - S - [KRHDE] - I - I - X - [GAVLIMPFW] - [GAVLIMPFW] -
 Q - S - R - C - X - X - F - R - F - X - P - [GAVLIMPFW]

Fig. 7:

Gleitklammerlader 2:

AC15_HUMAN : KAALLSGPPGVGKTTTASLV
MJ0884 : KPILLVGPPGCGKTTLAYAL
AF1195 : KPLLLAGPPGVGKTSALAL
MTH0240 : KPLLLVGPPGTGKTTLAHII
PHBN013 : KALLLAGPPGSGKTTTVYAL

Konsensus : KXXLLXGPPGXGKTSX\$XX\$

Pattern : K-X-X-L-L-X-G-P-P-G-X-G-K-T- [STNQYC] -X-
[GAVLIMPFW] -X-X- [GAVLIMPFW]Fig. 8:

Elongationprotein 1

DPOD_HUMAN : DVITGYNIQNFDLPYLISRA
DPOL_ARCFU : DIIVGYNQDAFDWPYLRKRA
MTH1208 : DILVGYNSDNFDFPYITRRA
PHBT047 : DVIITYNGDNFDFPYLLKRA
MJ0885 : DVIITYNGDNFDFPYLKARA

Konsensus : D\$XXYNXXXFDXPY\$XXRA

Pattern : D- [GAVLIMPFW] - [GAVLIMPFW] -X-X-Y-N-X-X-X-F-D-
X-P-Y- [GAVLIMPFW] -X-X-R-A

Fig. 9:

7/28

Elongationsprotein 2

AF1722 : AVRTAVAIMTEGVVAAPIEGIARVRI
 MJ1630 : AVRTALAVLTEGIVAAPLEGIADVKI
 PfuORF3 : AVRTALAILTEGIVSAPLEGIADVKI
 MTH1536 : ALRTALAILTEGVVAAPLEGIARVRI
 PHBN021 : AVRTALAILTEGVVSAPIEGIASVKI

Konsensus: A\$RTASA\$STEGSVXAP\$EGIAXV&I

Pattern : A- [GAVLIMPFW] -R-T-A- [GAVLIMPFW] -A- [GAVLIMPFW] -
 [GAVLIMPFW] -T-E-G- [GAVLIMPFW] -V-X-A-P- [GAVLIMPFW] -E-G-I-A-X-V-
 [KRHDE] -I

Fig. 10:

Elongationsprotein (Eubakteria)

DP3A_ECOLI : VPVGPGRGSGAGSLVAYALKITDLDPLEFDLLFERFLNPERVSMPD
 DP3A_SALTY : VPVGPGRGSGAGSLVAYALKITDLDPLEFDLLFERFLNPERVSMPD
 BB0579 : IPVGAGRGSGAGSIVAYALRITDIDPLKYNLLFERFLNPERISMPD
 DP3A_HELPY : IPVGPGRGSAAGSLVAFALKITDIDPLKYDLLFERFLNPERISMPD
 AA50 : IPVGPGRGSAGGSLVAYAIGITDVPDIKHGFLFERFLNPERVSMPD

Konsensus : \$PVG\$GRGSX\$GSS\$VAXA\$XITD\$DP\$XXX\$SLFERFLNPER\$SMPD

Pattern : [GAVLIMPFW] -P-V-G- [GAVLIMPFW] -G-R-G-S-X- [GAVLIMPFW] -
 G-S- [GAVLIMPFW] -V-A-X-A- [GAVLIMPFW] -X-I-T-D- [GAVLIMPFW] -D-P-
 [GAVLIMPFW] -X-X-X- [GAVLIMPFW] -L-F-E-R-F-L-N-P-E-R- [GAVLIMPFW] -S-M-
 P-D

Fig. 11:

Gleitklammer (Eubakteria) Region 1:

AAPOL3B : KLVITGGKSTYKLPTAPAEDFP
 DP3B_ECOLI : RMLVRSGRSRFSLSLTPAADFP
 S.TYPHIM. : RMLVRSGRSRFSLSLTPAADFP
 DP3B_PROMI : RLLVRSGRSRFSLSLTPASDFP
 DP3B_PSEPU : KLLVKAGRSRFTLSTLPANDFP
 DP3B_STRCO : RATVVCSSRFTLHTLPVEEYP

Konsensus : &XX\$XXGX\$XXXLXT\$P\$X&XP

Pattern : [KRHDE] -X-X- [GAVLIMPFW] -X-X-G-X-S-X-X-X-L-X-T-
 [GAVLIMPFW] -P- [GAVLIMPFW] -X- [KRHDE] -X-P

Gleitklammer (Eubakteria) Region 2:

AAPOL3B : IIMPMRV
 DP3B_ECOLI : VVMPMRL
 S.TYPHIM. : VVMPMRL
 DP3B_PROMI : VVMPMRL
 DP3B_PSEPU : VVMPMRL
 DP3B_STRCO : LIMPVRL

Konsensus : \$\$MP\$R\$

Pattern : [GAVLIMPFW] - [GAVLIMPFW] -M-P- [GAVLIMPFW] -R- [GAVLIMPFW]

9/28

Fig. 12:

```

      *           20           *           40           *
PCNA_HUMAN : LINFACWDISSGVNLQSMDSHVSQVLTSLRSEGFDTYRCDRNLAMGVNLTSMKIL : 58
PCNA_METJA : LLDEICFEVDEEGIKASAMDPHVALVSLEIPRLAFEEYEADS-HDIGIDLEAFKKVM : 57
MTH1312 : IVDEVQIQLSAEGLRLDALDRSHITYVHLELKAELEFDEYVCDEPERINVDTEELMKVL : 58
PHLA008 : LIDEAAFKVTEEGISMRAMPSPRVVLIDLNLPSISFSKYEVDEETIGVMNDHLKKVL : 58
AF0335 : LVSEARIHFLEKGLHSRAVDPANVAMVIVDIPKDSFEVYNIDEKTIQVMDRIFDIS : 58

      60           *           80           *           100           *
PCNA_HUMAN : KCAGNEDIITLRAEDNADTLALVFEAPNQEKVSDYEMKMLMDLVEQLG--IPEQEYSC : 114
PCNA_METJA : NRAKAKDRLILELDEEKNKLVIFENTG---KRKFSALLDISASSVK--VPEIEYPN : 110
MTH1312 : KRAKANDRVILSTDEG-N-LIIQFEGEA---VRTFKIRLIDIEYETPS--PPEIEYEN : 109
PHLA008 : KRGKAKDTLILRKGEENFLEISLQGT---TRTFRLPLIDVEEIEVE--LPDLPYTA : 110
AF0335 : KSISTKDLVELIVEDE-STLKVKFGSVE-----YKVALIDPSAIRKEPRIPELEPA : 109

      120           *           140           *           160           *
PCNA_HUMAN : VVKMPSGEFARICRDLSHIGDAVVISCAKDGVKFSASGELGNGNIKLSQTSNVKDKEE : 172
PCNA_METJA : VIMIKGDAFKEALKDADLFSQVILKADEKDFVIHAKGDLNENEAIFEKDSSA----- : 163
MTH1312 : EFEVPPFQLLKDSIADIDIFSDKITFRVDEDRFIASAEGEFGDAQIEY----- : 156
PHLA008 : KVVVLGEVLKEAVKDASLVSDSIKFMAKENEFIMRAEGETQEVVKLTLEDEG----- : 163
AF0335 : KIVMDAGEFKAIAAADKISDQVIFRSQKEGFRIEAKGDVDSIVFHMT-ET----- : 159

      180           *           200           *           220           *
PCNA_HUMAN : AVTIEMNEPVQVLTALRYLNFFTKATPLSSTVTLSMSADVPLVVEYKIADM-GHLKYY : 229
PCNA_METJA : IISLEVKEEAKSAFNLDYLMVMKGVSSGDIKIYLGNDMPLKLEYSIAG--VNLTF : 219
MTH1312 : LHGERIDKPARSIYSLDKIKEMLKADKFSETAIINLGDDMPLKLTLMASKEGELSFL : 214
PHLA008 : LLDIEVQEETKSAYGVSYLADMVKGIGKADEVMTMRFGNEMPMQMEYYIRDE-GRLTF : 220
AF0335 : ELIEFNGGEARSMFSVDYLKEFCVKVAGSGDLLTIHLGTNYPVRLVLFELVGGRAKVEYI : 217

PCNA_HUMAN : LAPKIED : 236
PCNA_METJA : LAPRIEG : 226
MTH1312 : LAPRIEA : 221
PHLA008 : LAPRVEE : 227
AF0335 : LAPRIES : 224

```

Fig. 13:

10/28

```

      *           20           *           40           *
AAPOL3B : MRVKVDREELEEVLLKKARESTEKKAALPILANFLLSAKEENLIVRATDLENYLVVSVK : 58
DP3B_ECOLI : MKFTVEREHLKPLQQVSGPLGGRTPLPILGNLLQVADGTLSTGTDLMEMVARVA : 58
S.TYPHIM. : MKFTVEREHLKPLQQVSGPLGGRTPLPILGNLLQVADGTLSTGTDLMEMVARVT : 58
DP3B_PROMI : MKFIIEREQLLKPLQQVSGPLGGRTPLPILGNLLKVNTLSTGTDLMEMMARVS : 58
DP3B_PSEPU : MHFTIQREALLKLQLVAGVVERQTLPVLSNVLLVVQGOQLSLTGTDLLEVELVGRVQ : 58
DP3B_STRCO : MKIRVERDVLAEAVAWAARSLPARPPAPVLAGLLLKAEEGQLSLSSFDYEV SARVSVE : 58

      60           *           80           *           100           *
AAPOL3B : GEVEEE-GEVCVHSQKLYDIVKNL-NSAYVYLHTEGEKLVITGGKSTYKLP TAPAEDF : 114
DP3B_ECOLI : LVQPHEPGATTVPARKFFDICRGLPEGAETAVQLEGERMLVRSGRSRFSLSTLPAADF : 116
S.TYPHIM. : LSQPHEPGATTVPARKFFDICRGLPEGAETAVQLEGERMLVRSGRSRFSLSTLPAADF : 116
DP3B_PROMI : LSQSHEIGATTVPARKFFDIWRGLPEGAETAVQLEGERMLVRSGRSRFSLSTLPAADF : 116
DP3B_PSEPU : LEEPAPGEITVPARKLMDICKSLPNDALIDIKVDEQKLLVKAGRSRFTLSTLPANDF : 116
DP3B_STRCO : AEIEEE-GTVLVSGRLLADISRAL-PNRPVEISTDGV RATVVCSSRFTLHTLPVEEY : 114

      120           *           140           *           160           *
AAPOL3B : PEFPEIVE-GGETLSGNLLVNGIEKVEYAIKKEANIALQGMYL RGYEDRIHFVGS DG : 171
DP3B_ECOLI : PNLDWQSEVEFTLPQATMKRLIEATQFSMAHQDVRYLNGMLFETEGEELRTVATDG : 174
S.TYPHIM. : PNLDWQSEVEFTLPQATMKRLIEATQFSMAHQDVRYLNGMLFETEGEELRTVATDG : 174
DP3B_PROMI : PNLDWQSEVEFTLPQATMKRLIEATQFSMAHQDVRYLNGMLFETEGEELRTVATDG : 174
DP3B_PSEPU : PTVEEGPGSLTCNLEQSKLRRLIERTSFAMAQDVRYLNGMLLEVSRNTLRAVSTDG : 174
DP3B_STRCO : PALPQMPE-ATGTVPGEVFASAVQQAIAAGRDDTL PVL TGVRIEIGDSVTLASTDR : 171

      180           *           200           *           220           *
AAPOL3B : HRLALYEPLGEFSKEL-----LIPRKS LKVLKCLITGIEDVNI EKSE---DESFA YFS : 221
DP3B_ECOLI : HRLAVCSMPIGQSLPS--HSVIVPRKGVIELMRMLDG-GDNPLRVQI---GSNNIRAH : 226
S.TYPHIM. : HRLAVCSMPLEASLPS--HSVIVPRKGVIELMRMLDG-GENPLRVQI---GSNNIRAH : 226
DP3B_PROMI : HRLAVCAMDIGQSLPG--HSVIVPRKGVIELMRMLDGSGESLLQLQI---GSNNIRAH : 227
DP3B_PSEPU : HRLALCSMSAPIEQEDR-HQVIVPRKGILELARLLTD-PEGMVSIVL---GOHHIRAT : 227
DP3B_STRCO : YRFAVREFLWKPENPDISAVALVP AKTLQDTAKALTSGDQVILALSGSGAGEGLIGFE : 229

      240           *           260           *           280           *
AAPOL3B : TPEWKLAVRLLEGEFPDYMSVIP EEFSAEVL FETEEVLKVLKRLKALSEGKVPVKIT : 279
DP3B_ECOLI : VGDFIFTSKLV DGRFPDYRRVLPKNPDKHLEAGCDLLKQAFARAAILSNEKFRGVRLY : 284
S.TYPHIM. : VGDFIFTSKLV DGRFPDYRRVLPKNPDKHLEAGCDILKQAFARAAILSNEKFRGVRLY : 284
DP3B_PROMI : VGDFIFTSKLV DGRFPDYRRVLPKNPDKHLEAGCDILKQAFARAAILSNEKFRGVRLY : 285
DP3B_PSEPU : TGEFTFTSKLV DGRFPDYRRVLPKGGDKLVVGDRLQALREAFSRTAILSNEKYRGIRLQ : 285
DP3B_STRCO : GAGRRTTTRLLEGDL PKYKTLFPTEFNSVAVIETAPFVEAVKRVALVA-ERNTPVRLS : 286

      300           *           320           *           340
AAPOL3B : LSENLAIFEADPEFGEAREEIEVEYTGEPFEIGFNGKYLMEALDAYDSERVWFKFTT : 337
DP3B_ECOLI : VSENQLKITANNPEQEEAEI LDVTSYSGAEMEIGFNVSYLDVLNALKCENVRMLT- : 341
S.TYPHIM. : VSENQLKITANNPEQEEAEI LDVTSYSGAEMEIGFNVSYLDVLNALKCETVRIMLT- : 341
DP3B_PROMI : LTNGQLKITANNPEQEEAEI LDVTSYSGAEMEIGFNVSYLDVLNALKCEVKKLLLT- : 342
DP3B_PSEPU : LAAGQLKIANNPEQEEAEI LDVTSYSGAEMEIGFNVSYLDVLNALKCEVKKLLLT- : 342
DP3B_STRCO : FEQGV LILEAGSSDDAQAVERVDAQLEGDDISIAFNPTFLDGLSAIDSPVAQLSFTT : 344

      *           360           *           380
AAPOL3B : ---PDTATLLEAEDYEK-EPYKCIIMP MRV-- : 363
DP3B_ECOLI : ---DSVSSVQIEDAAS-QSAAYVVM PMRL-- : 366
S.TYPHIM. : ---DSVSSVQIEDAAS-QSAAYVVM PMRL-- : 366
DP3B_PROMI : ---DAVSSVQVENVAS-AAAAVVM PMRL-- : 367
DP3B_PSEPU : ---DSNSSALLQEAGN-DDSSYVVM PMRL-- : 367
DP3B_STRCO : STKPALLSGRPAVDAEAEAYKYLIMPVRLSG ; 376

```

11/28

Fig. 14:

	*	20	*	40	*	60	
AC11_HUMAN	:	ICNYLSKIIPALQSRCTRFRFGPLTPELMVPRLEHVVEEEKVDISE	GMKALVTLS	SGDM	:	60	
AF2060	:	SCNYVSRIIEPIQSRCAVFRFKVPVKEAMKKRLLEICEKEGVKITED	GLEALIYIS	GGDF	:	60	
MTH0241	:	SCNYSSKIIDPIQSRCAIFRFLPLKGHQI IKRLEYIAEKENLEYEAHA	LETIVYFA	EAGDL	:	60	
PHBN012	:	SCNYSSKIIEPIQSRCAIFRFRPLRDEDI AKRLRYIAENEGLELTEE	GLQAILYA	EAGDM	:	60	
MJ1422	:	NCNYPISKIIPPIQSRCAVFRFSP LKKEDIAKKLKEIAEKEGLNL	TESGLEAI	IYVSEGDM	:	60	

	*	80	*	100	*	120	
AC11_HUMAN	:	RRALNILQSTNMAFGKVTEETVYTCTGHPLKSDIANILDWMLNQDFT	TAYRNITEL	KT	TLK	:	120
AF2060	:	RKAINALQGAAAIGEVVDADTIYQITATARPEEMTELIQTALKGNFME	ARELLDRL	MVEY	:	120	
MTH0241	:	RKAINLLQSAASLGEKITESSIYDVVSRARPKDVRKMIKTILDGKFME	ARDMLREI	MVLQ	:	120	
PHBN012	:	RRAINILQAAAALDKKITDENVMVASRARPEDIREMMLLALKGNFLK	AREKLR	EILLKQ	:	120	
MJ1422	:	RKAINVLQTAAALSDVIDDEIVYKVSSRARPEEVKKMMELALD	GKFMEAR	DLLYKLMVEW	:	120	

	*					
AC11_HUMAN	:	GLALHDILTEIHLFVHRVD	:	139		
AF2060	:	GMSGEDIVAQLFREIISMP	:	139		
MTH0241	:	GISGEDMVTQIYQELSRLA	:	139		
PHBN012	:	GLSGEDVLIQMHKEVFNLP	:	139		
MJ1422	:	GMSGEDILNQMFREINSLD	:	139		

Fig. 15:

12/28

```

      *           20           *           40           *
AC15_HUMAN : LLWVDKYKPTSLKTIIGQOGDQSCANKLLRWLRNWQKSSSEDKKHAAKFGKFSGKDDG : 58
MJ0884      : LSWVEKYRPKSLKDVAG---HEKVKEKLTWIESYLKGET----- : 37
AF1195      : MLWVEKYRPKTLEEVVA---DKSIITRVIKWAWSWKR----- : 35
MTH0240     : MSWTEKYRPGSFDEVVG---NQKVIAEIKEWIKAWKAGKP----- : 37
PHBN013     : VPWIEKYRPRKLSEIVN---QEQALEKVRWIESWLHGNPPK----- : 39

      60           *           80           *           100          *
AC15_HUMAN : SSFKAALLSGPPGVGKTTTASLVCQELGYSVELNASDTRSKSSLKAIVAESLNNTSI : 116
MJ0884      : --PKPILLVGPPGCGKTTLAYALANDYGFEVIELNASDKRNSSAIKKVVGHAATSSSI : 93
AF1195      : --SKPLLAGPPGVGKTSALALANTMGWEAVELNASDQRSWRVIERIVGEGAFNETI : 91
MTH0240     : --QKPLLVGPPGTGKTTLAHIIGKEFS-DTLELNASDRRSQDALMRSAGEASATRSL : 92
PHBN013     : --KKALLLAGPPGSGKTTTVALAHEYNFVIELNASDERTYNKIARYV-QAAYTMDI : 94

      120          *           140          *           160          *
AC15_HUMAN : KGFYSNGAASSVSTKHALIMDEVDGMAGNEDRGGIQELIGLIK-HTKIPIICMCNDRN : 173
MJ0884      : FG-----KKF-LIVLDEVDGISGKEDAGGVSELIKVIK-KAKNPIILTANDAY : 139
AF1195      : SDEGEF-LSSRIGKLLIILDEVNDIHKKEDVGGEAALIRLIKRPQAQPLILIANDPY : 148
MTH0240     : FN-----HDLKLIILDEVDDGIHGNEDRGGVQAINRIK-ESRHPMVLITANDPY : 139
PHBN013     : MG-----KRRKIIFLDEADNIE--P--SGAPEIAKLID-KARNPIIAMAANYH : 137

      180          *           200          *           220          *
AC15_HUMAN : HPKIRSLVHYCFDLRFQRPVEQIKGAMMSIAFKEGLKIPPPAMNEIILGANQDIRQV : 231
MJ0884      : APSIRSLPYVEVIQLNPVHTNSVYKVLKKIAEKEGLDVDDKTLKMAIQAHSAGDLRSA : 197
AF1195      : KLSPELRNLCEMINFKRLTKQVARVLERIALKEGIKVDKSVLLKIAENAGGDLRAA : 205
MTH0240     : SKRLQSIKPRCRVLNLRKVHTSSIAAALRRICRAEGIECPDDVLRRELAKRSRDLRSA : 197
PHBN013     : EVPK-EIRDRAELVEYKRLNQRDVISALVRILKREGITVPKEILTEIAKRSSGDLRAA : 194

      240          *           260          *           280          *
AC15_HUMAN : LHNLSMWCARSKALTYDQAKADSHRAKDIKMGPFVDVARKVFAAGEETAHMSLVDKSD : 289
MJ0884      : INDLEALALSGDLSYEAQKLP----DRKREANIFDALRVILKTTTHYGIATTALMN-- : 249
AF1195      : INDQALAEKGKELKPEDVFLT---KRTQEKDIFRVMQMIKTKNPAVYNEAML--- : 256
MTH0240     : INDLEAMAEGEERIGEELLKMG---EKDATSNLFDVAVRLKSRDVSQVREAMR--- : 248
PHBN013     : INDLQTTIVAGG--YEDAKYVLA---YRDVEKTVFQSLGMVFSSDNAKRAKLALMN-- : 244

      300          *           320          *           340
AC15_HUMAN : LFFHDYSIAPLFVQENYIHVKPVAAGGDMKKHMLLSRAADSICDGLVDSQIRSKQN : 347
MJ0884      : -----VDETPDVVIEWIAENVPKKEYEKPEEVARAFEYLSKADRYLGRVMRRQN--YSF : 300
AF1195      : -----LDESPEDVIHWVDENLPLEYSG-VELVNAYEALSRAIDFLGRVRRRQF--YRL : 306
MTH0240     : -----VDDDPVLVLEFIAENVPREYKPNISRAYDMLSRADIFFGRAVTRN--YTY : 299
PHBN013     : -----LDMSPDEFLLWVDENIPHYLKPPEEMARAYEASIRADIYLGRAQRTGN--YSL : 295

      *           360          *           380          *           400
AC15_HUMAN : WSLPQAQAIYASVLPGLMRGYMTQFPFPWSWLGKHSSTGKHDRIVQDLALHMSLRTY : 405
MJ0884      : WKYATTLMTAGVALSKDEKYRKWTPYSY-PKIFRLLTKTKAEREILNKLKIGEKTH : 357
AF1195      : WKYASYLMTVGVOQMKEEPPKGFTRYRR-PAVWQMLFQLRQKREMTKILEKIGKYSH : 363
MTH0240     : WRYASELMGPGVALAKDKTYRKVRYTG-SSSFRLGKTRKQSRSDSVAAKMAGKMH : 356
PHBN013     : WKYAIMMTAGVAVAGTK-KKGFAYFYP-PNTLKMLESKEERSIRDSIIKKIMKEMH : 351

      *           420
AC15_HUMAN : SSKRTVNMDYLSLLR : 420
MJ0884      : TSSKRAR-FDLQMLK : 371
AF1195      : LSMRKARTEMFPVIK : 378
MTH0240     : ISPKVAI-SMFPYME : 370
PHBN013     : MSKLEAL-ETMKILR : 365

```

Fig. 16:

```

      *           20           *           40           *
PHBN023 : IFEVEDQTRVKVFLPKDSED-YREALKVLPAVVAFAKGVYSKRG-IFFANRFYLPDV : 56
PfuORF2 : IFEIEDLTGKVKVFLPKDSED-YREAFKVLPAVVAFAKGVYSKRG-ILYANKFYLPDV : 56
DPD2_HUMAN : -LVLEDELQRIKLGKTIDVSK-----LVTGTVLAVFGSVRDDGKFLVEDYCFADLA : 50
MTH1405 : IIELEDDTGEISVVVHNENHKLFEKSEKIVRDEVVGVHGTKKGR--FVVASEIFHPGV : 56
AF1790 : YIRLEDTTGTITCVATGKNAE---VARELLGDEVIGVTGLLKGS--SLYANRIVFPDV : 53
MJ0702 : IVRIEDTEDEATLILPKIEAGKIPDDILLDEVIGAIGTVSKSGSSIYVDEIRPAL : 58

      60           *           80           *           100           *
PHBN023 : PLYRK-QKPPLEEKVYAVLTSDIHVGSK--EFCEKAFIKFLEWLNGYVESKEEEEIVS : 111
PfuORF2 : PLYRR-QKPPLEEKVYAILISDIHVGSK--EFCENAFIKFLEWLNGNVETKEEEEIVS : 111
DPD2_HUMAN : PQKP--APP-LDTRFVLLVSGGLGGGGGESLLGT-QLLVDDVVTGQLGDEGEQCSAA : 104
MTH1405 : PRIQ--EK---EMDFSVAFISDVHIGSQ--TFLEDAFMKFVKWINGDFGSEEQSRSLAA : 107
AF1790 : PINGNGEK---KRDFYIVFLSDTHFGSK--EFLEKEWEMFVRWLKGEVGGKKSQNLAE : 106
MJ0702 : PPKEP-KR--IDEEIYMAFLSDIHVGSK--EFLHKEFEKFIKFLNGDVEDNEEEKVVS : 111

      120           *           140           *           160           *
PHBN023 : RIRYLIAGDVVD-GIGIYP-GQYSDLIIPDIFDQYEALANLLSNVP----KHITIFI : 163
PfuORF2 : RVKYLIAGDVVD-GVGVPY-GQYADLTIPDIFDQYEALANLLSHVP----KHITMFI : 163
DPD2_HUMAN : HVSRLVILAGNLLSHSTQSRDSINKAKYLTKKTQAASVEAVKMLDEILLQLSASVPVDV : 162
MTH1405 : DVKYLVVAGDIVD-GIGIYP-GQEKELLIRDIHEQYEEAARLFGDIR----SDIKIVM : 159
AF1790 : KVKYIYIAGDIVD-GIGVYP-GQEDDLAISDIYQYEFASHLDEIP----KEIKIIV : 158
MJ0702 : RLKYICIAGDLVD-GVGVPY-GQEDLYEVDIIEQYREIAMYLDQIP----EHISIII : 163

      180           *           200           *           220           *
PHBN023 : GPGNHDAARPAIPQPEFYEEYAKPLYKLKNTVVIISNPAVIRLHGRDFLIAHGRGIEDV : 221
PfuORF2 : APGNHDAARQAIPQPEFYKEYAKPIYKLKNAVIISNPAVIRLHGRDFLIAHGRGIEDV : 221
DPD2_HUMAN : MPGEFDPNTYTLPPQPLHPCMFPLATAYSTLQLVNPNYQATIDGVRFLGTSGQNVSDI : 220
MTH1405 : IPGNHDSSRIAEPQPAIPEEYAKSLYSIRNIEFLSNPSLVSLDGVRTLIYHGSRFDDM : 217
AF1790 : SPGNHDAVRQAEPQPAFEGEI-RSLFP-KNVEHVGNPAYVDIEGVKVLIIYHGSRIDDI : 214
MJ0702 : SPGNHDAVRPAEPQPKLPEKITKLFNR-DNIYFVGNPCTLNHGFDTLLYHGSRFDDL : 220

      240           *           260           *           280           *
PHBN023 : VSFVPGLTHHKPGLPMVELLKMRLHAPTFGGKVPIAPDPE-DLLVIEEVPDLVQMGHV : 278
PfuORF2 : VGSVPGLTHHKPGLPMVELLKMRLHVAPMFGGKVPIAPDPE-DLLVIEEVPDVVHMGHV : 278
DPD2_HUMAN : FRYSG---SMEDHLEILEWTLRVRHISPTAPDTLGCYPFYKIDPFIPECPHVYFCGNT : 275
MTH1405 : AMSVNGLSHERSDLIMEELLEKRHLAPIYGERTPLASEIE-DHLVIDEVPVHLHTGHV : 274
AF1790 : ISKIPRLSYDEPQKVMELLKRRHLSPIYGGRTPLAPERE-DYLVIEDVPDILHCGHI : 271
MJ0702 : VGQIRAASYENPVTIMKELIKRRLLCPTYGGRCPIAPEHK-DYLVIDRDIDILHTGHI : 277

      300           *           320           *
PHBN023 : HVYDTAVYRG-----VQLVNSATWQAQTEFQKMVNIVPTPGLVPIVD : 320
PfuORF2 : HVYDAVVYRG-----VQLVNSATWQAQTEFQKMVNIVPTPAKVPVVD : 320
DPD2_HUMAN : PSFGSKIIRGPEDQTVLLVTVPDF-SATQTAACLVNLR-SLACQPISF : 320
MTH1405 : HINAYKKYKG-----VHLINSGTFQSQTEFQKIYNIVPTCGQVPVLN : 316
AF1790 : HTYGTGFYRG-----VFMVNSSTWQAQTEFQKKVNLNPMPGNVAVYR : 313
MJ0702 : HINGYGIYRG-----VVMVNSGTFQEQTDFQKRMGISPTPAIVPIIN : 319

```

14/28

Fig. 17:

```

          *      20      *      40      *
DPOD_HUMAN : KVQSYEKEEDLLQAWSTFIRIMDPDVTIGYNIQNFDLPYLISRAQTLKVQTFPFLGRV : 58
DPOL_ARCFU  : EIILTGDERKIIISDFVKLIKSYDPDIIVGYNQDAFDWPYLRKRAERWNIPLD--VG-- : 54
MTH1208     : FVEVVEDERELLERFAEIVIDKKPDILVGYNSDNFDFFPYITRRAAILGAELD--LG-- : 54
PHBT047     : YVEVVSSEMERIKRLIRVIKEKDPDVIITYNGDNFDFPYLLKRAEKLGIKLL--LG-- : 54
MJ0885      : NIEVVKNEKELIKKIIETLKEY--DVIITYNGDNFDFPYLKARAKIYGIDIN--LG-- : 52

          60      *      80      *      100      *
DPOD_HUMAN : AGLCSNIRDSSFQSKQTGRRDTKVSMVGRVQMDMLQVLLREYKLSHTLNAVSVFHL : 116
DPOL_ARCFU  : -----RDGSNVVFRGG-----RPKITGRNLVDLYDIAMRISDIKIKKLENVAEFLG : 100
MTH1208     : -----WDGSKIRTMRRG-FANATAIKGTVHVDLYPVMRRYMNLDRYTLERVYQELF : 104
PHBT047     : -----RDNSEPKMQKMG-DSLAVEIKGRIHFDLPVIRRTINLPTYTLEAVYEAIF : 104
MJ0885      : -----KDGEELKIKRGG-MEYRSYIPGRVHIDLYPISRRLKLTKYTLEDVVYNLF : 102

          120      *      140      *      160      *
DPOD_HUMAN : GEQKE-DVQHSIITDLQNGNDQTRRRRLAVYCLKDAYLPLRLLERLMVLVNAVEMARVT : 173
DPOL_ARCFU  : TKIEIADIEAKDIYRYWSRGE--KEKVLNYARQDAINTYLIK--ELLPMHYELSKMI : 154
MTH1208     : GEEKI-DLPGDRLWEYWDREDEL-RDELFRYSLDDVVATHRIAE--KILPLNLELTRLV : 158
PHBT047     : GKPKE-KVYADEIAKAWETGEG-LERVAKYSMEDAKVTYELGR--EFFPMEAQLARLV : 158
MJ0885      : GIEKL-KIPHTKIVDYWANND---KTLEIYSLQDAKYTYKIGK--YFFPLEVMFSRIV : 154

          180      *      200      *      220      *
DPOD_HUMAN : GVPLSYLLSRGQQVKVVSQLLRQAMHEGLLMPVVKSE-----GGEDYTGATVIEPLK : 225
DPOL_ARCFU  : RLPVDDVTRMGRGKQVDWLLLSEAKKIGEIAPNPPE-----HAESYEGAFVLEPER : 205
MTH1208     : GQPLFDISRMATGQQAWEFLVRKAYQYGELVPNKPSQSDFFSSRRGRRVGGYVKEPEK : 216
PHBT047     : GQPVWDVSRSSGTNLVEWFLLRKAYERNELAPNKPDEKEYERRRLRESYEGGYVKEPEK : 216
MJ0885      : NQTFPEITRMSSGQMVEYLLMKRAFKENMIVPNKPDEEYRRRVLTITYEGGYVKEPEK : 212

          240
DPOD_HUMAN : GYYDVPIATLDFS : 238
DPOL_ARCFU  : GLHEN-VACLDFA : 217
MTH1208     : GLHEN-IVQFDFR : 228
PHBT047     : GLWEG-IVSLDFR : 228
MJ0885      : GMFED-IISMDFR : 224

```

Fig. 18:

```

      *           20           *           40           *           60
AF1722 : DTIKGVKGMTSKTKIPERLEKGILRVKHGVFVKDGTARFDTATDLPITHFKPAEIGVSVEKLR : 63
MJ1630 : GDVKCIKGMTSKQKIVEPLEKAILRAINEVYVFKDGTTRFDCTDVPVTHFKPNEINVTVEKLR : 63
PfuORF3 : DKLGVMGMTSGWKIAEPELEKGLLRKANEVYVFKDGTIRFDATDAPITHFRPREIGVSVEKLR : 63
MTH1536 : DEIKGVEGMISA EKFPPELEKGLLRKANDVYTFKDATIRHDS TDLP LTHFTPREVGVSVERLR : 63
PHBN021 : DKLGVMGMTSGWKMPPELEKGLLRKANDVYVFKDGTIRFDATDAPITHFRPREIGVSVEKLR : 63

      *           80           *           100          *           120
AF1722 : ELGYERDYKGAELKNENQIVELKPQDVILPKSGAEYLLRVANFIDDLVKFYKMEPFYNAKSV : 126
MJ1630 : ELGYDKDIYGNELVDGEQVVELKPQDVIIIPESCAEYFVKVANFIDDLLEKFKYKVERFYNVKKK : 126
PfuORF3 : ELGYTHDFEGKPLVSEDQIVELKPQDVILSKEAGKYLLRVARFVDDLLEKFGYGLPRFYNAEKM : 126
MTH1536 : ELGYTRDCYGDELEDEDQILELRVQDVVISED CADYLVRVANFVDDLLERFYDLERFYNVKTR : 126
PHBN021 : ELGYTHDFEGNPLVSEDQIVELKPQDIILSKEAGKYLLKVAKFVDDLLEKFGYGLPRFYNAEKM : 126

      *           140          *           160           *
AF1722 : EDLIGHLVIGLAPHTSAGVLGRIIGFSDVLAGYAHYPYFHAARR : 170
MJ1630 : EDLIGHLVIGMAPHTSAGMVGRIIGYTKANVGYAHYPYFHAARR : 170
PfuORF3 : EDLIGHLVIGLAPHTSAGIVGRIIGFVDALVGYAHYPYFHAARR : 170
MTH1536 : EDLVGHLIAGLAPHTSAAVLGRIIGFTGASACYAHYPYFHSARR : 170
PHBN021 : EDLIGHLVIGLAPHTSAGIVGRIIGFVDALVGYAHYPYFHAARR : 170

```

Fig. 19:

```

      *           20           *           40           *           60
DP3A_ECOLI : ELQVINQMGPFGYFLIVMEFIQWSKDNQVPGPGRGSGAGSLVAYALKITDLDPLEFDLL : 60
DP3A_SALTY : ELQVINQMGPFGYFLIVMEFIQWSKDNQVPGPGRGSGAGSLVAYALKITDLDPLEFDLL : 60
BB0579 : ELSVIIGMGFEGYFLIVWDFIKFAHDNDIPVGAGRGSGAGSIVAYALRITDIDPLKYNLL : 60
DP3A_HELPY : EIEVITNMKFPGYMLIVWDFIRYAKEMGIPVGPGRGSAAGSLVAFALKITDIDPLKYDLL : 60
AA50 : ELEVINMGFAGYFLIVQDFINWAKKNDIPVGPGRGSGAGSLVAYAIGITDVPDKHGFL : 60

      *           80           *           100          *           120
DP3A_ECOLI : FERFLNPERVSMDFDVFDFCMEKRDQVIEHVADMYGRDAVSQIITFGTMAAKAVIRDVGR : 120
DP3A_SALTY : FERFLNPERVSMDFDVFDFCMEKRDQVIEHVADMYGRDAVSQIITFGTMAAKAVIRDVGR : 120
BB0579 : FERFLNPERISMDFDIDDFCFEGRDEIIKYVTNKYGEDKVAQIITFGTLKPKAVKDVVAR : 120
DP3A_HELPY : FERFLNPERISMDFDIDTDFCQRRRKEIIEYMIKYGKYNVAQVITFNKMLAKGVIRDVAR : 120
AA50 : FERFLNPERVSMDFDIDVFCDNREKVIYVRNKYGHNDVAQIITYNVMAKQTLRDVAR : 120

      *
DP3A_ECOLI : VLGHYPYGFVDRISKLIIP : 138
DP3A_SALTY : VLGHYPYGFVDRISKLVPP : 138
BB0579 : VLDIPFAESNELTKFIPD : 138
DP3A_HELPY : VLDMPYKEADDFAKLIPN : 138
AA50 : AMGLPYSTADKLAKLIPQ : 138

```

Fig. 20

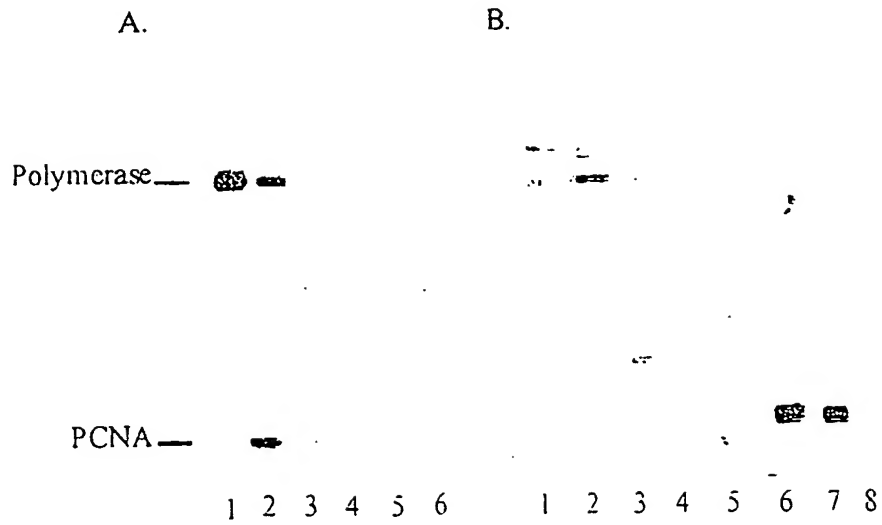


Fig. 21

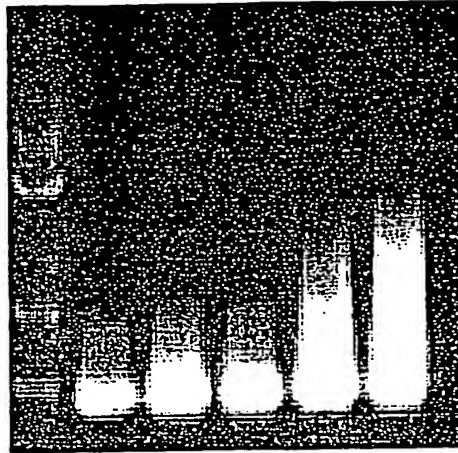
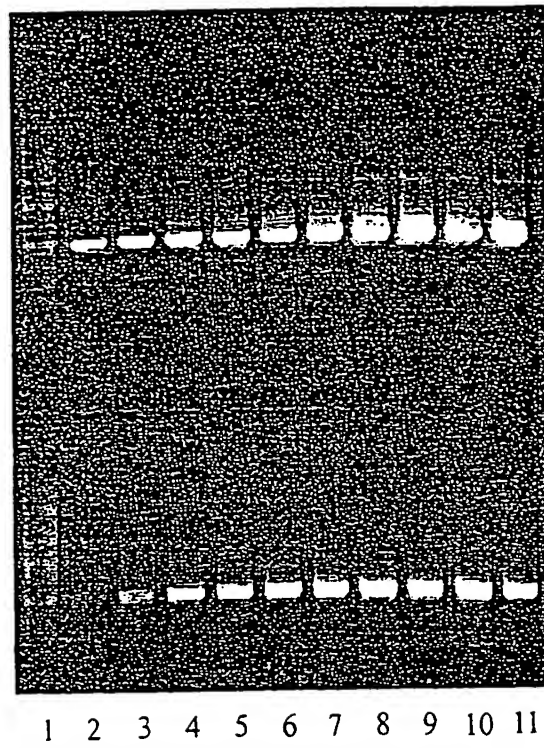


Fig. 22



19/28

Fig. 23

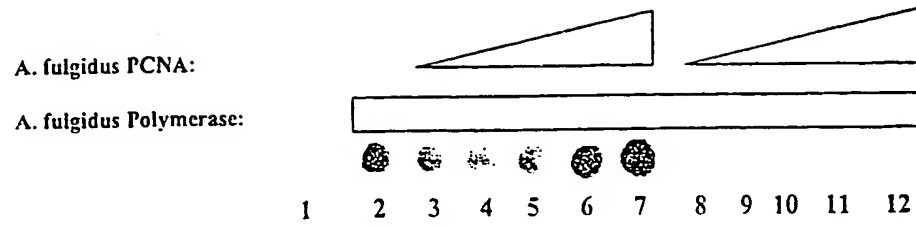


Fig. 24

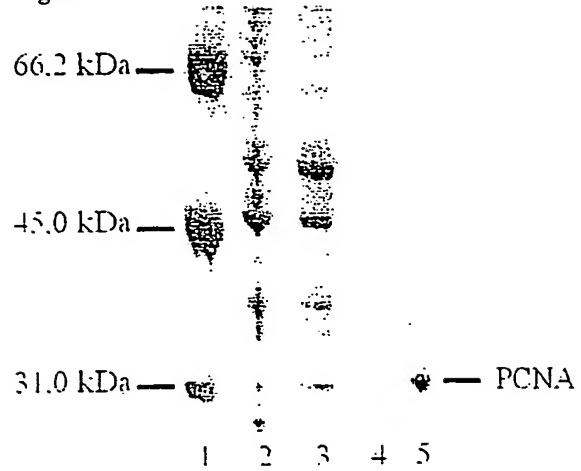


Fig. 25

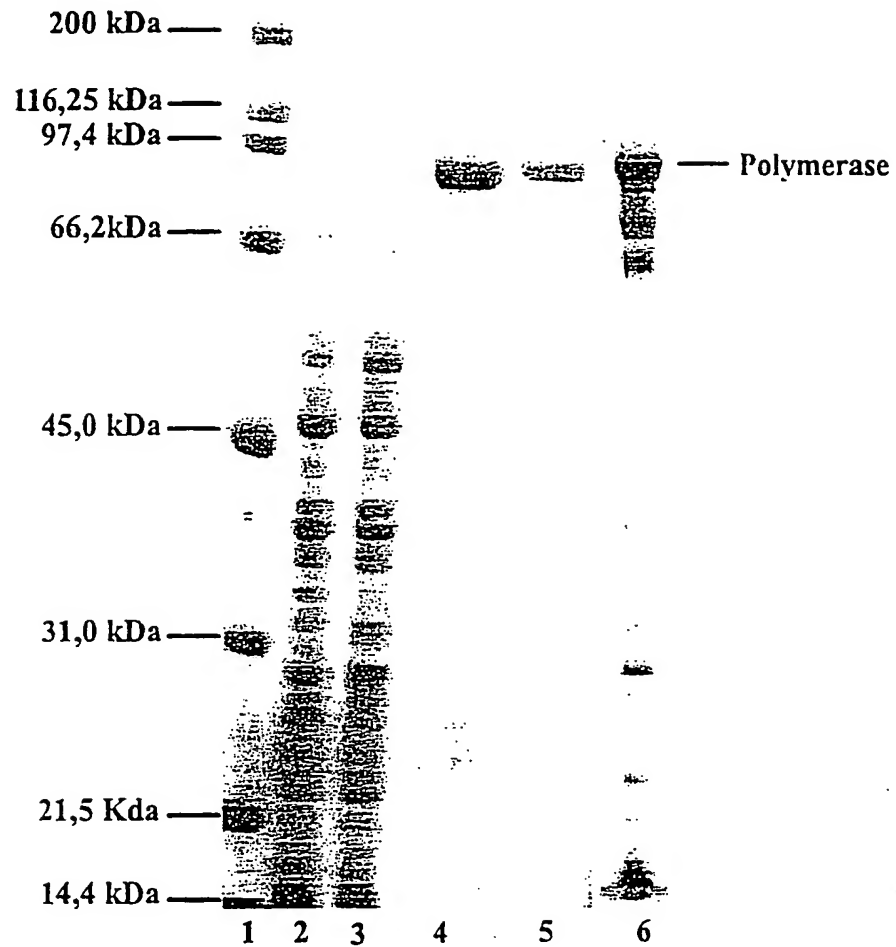


Fig. 26



1 2

Fig. 27

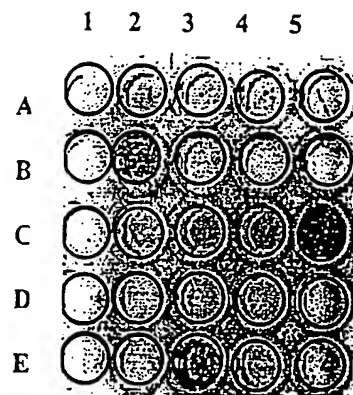


Fig. 28

1 2 3

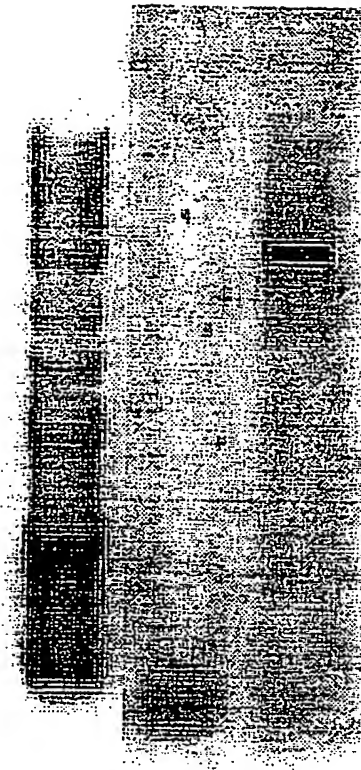


Fig. 29 a

Name des Gens (wie SRS gefunden)	Genname (wie in Fig. 1 angegeben)	Zugangs- nummer	Datenbank	Zugangsnummer	Datenbank	Seq ID NO
AC11 HUMAN	Identisch	L07540	EMBL	P40937	SWISSPROT	1
AC12 HUMAN	Identisch	M87339	EMBL	P35249	SWISSPROT	32
AC13 HUMAN	Identisch	L07541	EMBL	P40938; O15252	SWISSPROT	33
AC14 HUMAN	Identisch	M87338	EMBL	P35250; P32846	SWISSPROT	34
AC15 HUMAN	Identisch	L14922	EMBL	P35251	SWISSPROT	6
DPD2 HUMAN	Identisch	U21090	EMBL	P49005	SWISSPROT	16
DPOD HUMAN	Identisch	M80397	EMBL	P28340	SWISSPROT	22
PCNA HUMAN	Identisch	M15796	EMBL	P12004	SWISSPROT	11
AF0335	Identisch	AE001081	TREMBL	O29912	SPTREMBL	12
AF0497	Identisch	AE001070	GENPEPT	O29753	SWISSPROT	23
AF1195	Identisch	AE001022	TREMBL	O29072	SPTREMBL	7
AF1722	Identisch	AE000984	TREMBL	O28552	SPTREMBL	27
AF1790	Identisch	AE000979	TREMBL	O28484	SPTREMBL	17
AF2060	Identisch	AE000961	TREMBL	O28219	SPTREMBL	2
MJ0247	Identisch	U67480	TREMBL	Q57697	SWISSPROT	13
MJ0702	Identisch	U67516	TREMBL	Q58113	SWISSPROT	18
MJ0884	Identisch	U67532	TREMBL	Q58294	SPTREMBL	8
MJ0885	Identisch	U67532	TREMBL	Q58295	SWISSPROT	24
MJ1422	Identisch	U67583	TREMBL	Q58817	SPTREMBL	3
MJ1630	Identisch	U67603	TREMBL	Q59024	SPTREMBL	28
MTH1208	Identisch	AE000888	TREMBLNEW	O27276	SWISSPROT	26
MTH1312	Identisch	AE000895	TREMBLNEW	O27367	SWISSPROT	15
MTH1405	Identisch	AE000903	GENPEPT	O27456	SPTREMBL	20
MTH1538	Identisch	AE000913	TREMBLNEW	O27579	SPTREMBL	30
MTH240	MTH0240	AE000811	TREMBLNEW	O26342	SPTREMBL	10
MTH241	MTH0241	AE000811	TREMBLNEW	O26343	SPTREMBL	5
PolC	PFUORF2	D84670	EMBL	P81409	SPTREMBL	21

Fig. 29 b

	PoIB	PFUORF3	D84670	EMBL	P81412	SPTREMBL	31
	PH0112	PHBN012	AP000001_116	TREMBL	O57852	SPTREMBL	4
	PH0113	PHBN013	AP000001_117	TREMBL	O57853	SPTREMBL	9
	PHLA008	PH008	AP000001	EMBL	O57772	SPTREMBL	14
	PH0123	PHBN023	AP000001	EMBL	O57863	SPTREMBL	19
	DPOL PYRHO	PHBT047	AP000007	EMBL	O59610	SWISSPROT	25
	PH0121	PHBN021	AP000001	TREMBL	O57861	SPTREMBL	29
	DP3B ECOLI	Identisch	K02179	EMBL	P00583	SWISSPROT	35
	DP3A ECOLI	Identisch	M19334	EMBL	P10443	SWISSPROT	37
	DNAN	AASEQ93	AE000761	EMBL	O67725	SWISSPROT	36
	DNAE	AASEQ50	AE000718	EMBL	O67125	SWISSPROT	38

Fig. 30a

EMBL	
URL	http://www.ebi.ac.uk
Status	Die aktuelle Version besteht aus 3952878 Einträgen und wurde am 30. Juli 1999 registriert.
Beschreibung	Das Europäische Bioinformatik-Institut (EBI) unterhält und vertreibt die 'EMBL Nucleotide Sequence database', Europas primäre Quelle für Nukleotidsequenzdaten. Das EBI unterhält und vertreibt in Zusammenarbeit mit Amos Bairoch von der Universität Genf ebenfalls die 'SWISS-Prot Protein Sequence database'. Über fünfzig zusätzliche, spezialisierte molekularbiologische Datenbanken sind verfügbar, ebenso wie Software und Dokumentationen, die für Molekularbiologen von Interesse sind. Die EBI Netzwerk-Dienste umfassen Datenbanksuche und Sequenzvergleichseinrichtungen.
Literatur	Patricia Rodríguez-Tomé, Peter J. Stoeckli, Graham N. Cameron und Tomas P. Flores, "The European Bioinformatics Institute (EBI) databases", <i>Nucleic Acids Res.</i> 24:76-13, 1996.

Fig. 30b

SWISSPROT	
URL	http://expasy.hcuge.ch/spmot/spmot.fon.html
Status	Die aktuelle Version besteht aus 80000 Einträgen und wurde am 30. Juli 1999 registriert
Beschreibung	Die 'SWISS-Prot Protein Sequence database' ist eine Datenbank mit Proteinsequenzen, die in Zusammenarbeit von Amos Bairoch (Universität Genf) mit der EMBL Daten-Bibliothek erstellt wurde. Die Daten in Swiss-Prot sind abgeleitet von DNA-Sequenz-Übersetzungen aus der 'EMBL Nucleotide Sequence Database', übernommen aus der PIR-Sammlung (Protein Identification Resource), der Literatur entnommen oder direkt von Forschern übermittelt worden. Sie enthält hochwertige Anmerkungen, ist nicht-redundant und bietet Querverweise zu diversen anderen Datenbanken, u.a. der 'EMBL nucleotide sequence database', der 'PROSITE pattern database' und zur PDB. SWISS-Prot ist eine gepflegte Proteinsequenzdatenbank, die bemüht ist, ein hohes Maß an Annotationen (wie die Beschreibung der Funktion eines Proteins, seiner Domänenstruktur, posttranslationale Modifikationen, Varianten, etc.), ein minimales Maß an Redundanz und ein hohes Maß an Integration mit anderen Datenbanken zu liefern. Die neuesten Entwicklungsstufen der Datenbank beinhalten: eine Erhöhung von Anzahl und Umfang von Modellorganismen, Querverweise zu sieben zusätzlichen Datenbanken, eine Vielzahl von neuen Dokumentationen, die Erstellung von TrEMBL, einem nicht-annotierten Zusatz zu Swiss-Prot. Dieser Zusatz besteht aus Einträgen in einem Swiss-Prot-ähnlichen Format, die von der Übersetzung aller kodierenden Sequenzen (CDS) der EMBL nucleotide sequence database abgeleitet sind, sofern die CDS nicht schon in Swiss-Prot enthalten sind.
Literatur	Amos Bairoch and Rolf Apweiler: "The SWISS-PROT protein sequence data bank and its supplement TREMBL", Nucleic Acids Res. 25:(31-36), 1997

Fig. 30c

TrEMBL	
URL	http://expasy.hcuge.ch/sprot/sprot-top.html
Status	Die aktuelle Version besteht aus 411044 Einträgen und wurde am 30. Juli 1999 registriert.
Beschreibung	TrEMBL ist eine Computer-annotierte Proteinsequenz-Datenbank, die ein Zusatz zur Proteinsequenzdatenbank SWISSPROT ist. TrEMBL enthält die Übersetzung aller codierender Sequenzen (CDS), die in der EMBL-Nukleotidsequenzdatenbank vorhanden und noch nicht in SWISSBROTT integriert sind. TrEMBL kann als temporäre Vorstufe zu Swiss-Prot betrachtet werden. Allen Einträgen in TrEMBL, die später zur Standard-Swiss-Prot-Qualität aufgewertet werden sollten, wurden Swiss-Prot-Zugangsnummern zugeteilt. TrEMBL ist in zwei Hauptsektionen eingeteilt: SPTREMBL und REMTREMBL. SPTREMBL (SWISSPROT TrEMBL) enthält die Einträge, die letztendlich in SWISSPROT eingebaut werden sollten. SWISSPROT-Zugangsnummern wurden für alle SP-TrEMBL-Einträge vergeben.
Literatur	Amos Bairoch und Rolf Apweiler "The SWISS-PROT protein sequence data bank and its supplement TrEMBL", <i>Nucleic Acids Res.</i> 25(31-35), 1997

STREMBL	
URL	http://www.ebi.ac.uk/~sp/information.html
Status	Die aktuelle Version besteht aus 199794 Einträgen und wurde am 4. August 1999 registriert.
Beschreibung	siehe TrEMBL
Literatur	siehe TrEMBL

TREMBLNEW	
URL	nicht verfügbar
Status	Die aktuelle Version besteht aus 71306 Einträgen und wurde am 4. August 1999 registriert.
Beschreibung	Diese Datenbank enthält die Updates zur letzten vollständigen Version von TrEMBL.
Literatur	siehe TrEMBL

GENPEPT	
URL	http://www.infobiogen.fr/page_accueil_en.html
Status	Die aktuelle Version besteht aus 412243 Einträgen und wurde am 30. Juli 1999 registriert.
Beschreibung	GENPEPT ist eine Protein-Datenbank, die eine Übersetzung der neuesten Version der GENBANK darstellt. Das Programm „Create_GenPept“ wurde von Mark A. Gunnell am NCI, USA geschrieben und von G. Vaysseix bei INFOBIOGEN, Frankreich überarbeitet, um auf modernen Computersystemen (SUN/UNIX) lauffähig zu sein.
Literatur	keine Quellen verfügbar

SEQUENZPROTOKOLL

(1) ALLGEMEINE ANGABEN:

(i) ANMELDER:

- (A) NAME: LION bioscience AG
- (B) STRASSE: Im Neuenheimer Feld 517
- (C) ORT: Heidelberg
- (E) LAND: DE
- (F) POSTLEITZAHL: 69120

(ii) BEZEICHNUNG DER ERFINDUNG: Akzessorische Komplexe mit Polymeraseaktivitaet

(iii) ANZAHL DER SEQUENZEN: 54

(iv) COMPUTER-LESBARE FASSUNG:

- (A) DATENTRÄGER: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) BETRIEBSSYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPA)

(2) ANGABEN ZU SEQ ID NO: 1:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 340 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: beides

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: Homo sapiens

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 1:

Met Glu Thr Ser Ala Leu Lys Gln Gln Glu Gln Pro Ala Ala Thr Lys
1 5 10 15

Ile Arg Asn Leu Pro Trp Val Glu Lys Tyr Arg Pro Gln Thr Leu Asn
20 25 30

Asp Leu Ile Ser His Gln Asp Ile Leu Ser Thr Ile Gln Lys Phe Ile
35 40 45

Asn Glu Asp Arg Leu Pro His Leu Leu Leu Tyr Gly Pro Pro Gly Thr
50 55 60

Gly Lys Thr Ser Thr Ile Leu Ala Cys Ala Lys Gln Leu Tyr Lys Asp
65 70 75 80

Lys Glu Phe Gly Ser Met Val Leu Glu Leu Asn Ala Ser Asp Asp Arg
85 90 95

Gly Ile Asp Ile Ile Arg Gly Pro Ile Leu Ser Phe Ala Ser Thr Arg
100 105 110

Thr Ile Phe Lys Lys Gly Phe Lys Leu Val Ile Leu Asp Glu Ala Asp
115 120 125

Ala Met Thr Gln Asp Ala Gln Asn Ala Leu Arg Arg Val Ile Glu Lys
130 135 140

Phe Thr Glu Asn Thr Arg Phe Cys Leu Ile Cys Asn Tyr Leu Ser Lys
145 150 155 160

Ile Ile Pro Ala Leu Gln Ser Arg Cys Thr Arg Phe Arg Phe Gly Pro
165 170 175

Leu Thr Pro Glu Leu Met Val Pro Arg Leu Glu His Val Val Glu Glu
180 185 190

Glu Lys Val Asp Ile Ser Glu Asp Gly Met Lys Ala Leu Val Thr Leu
195 200 205

Ser Ser Gly Asp Met Arg Arg Ala Leu Asn Ile Leu Gln Ser Thr Asn
210 215 220

Met Ala Phe Gly Lys Val Thr Glu Glu Thr Val Tyr Thr Cys Thr Gly
225 230 235 240

His Pro Leu Lys Ser Asp Ile Ala Asn Ile Leu Asp Trp Met Leu Asn
245 250 255

Gln Asp Phe Thr Thr Ala Tyr Arg Asn Ile Thr Glu Leu Lys Thr Leu
260 265 270

Lys Gly Leu Ala Leu His Asp Ile Leu Thr Glu Ile His Leu Phe Val
275 280 285

His Arg Val Asp Phe Pro Ser Ser Val Arg Ile His Leu Leu Thr Lys
290 295 300

Met Ala Asp Ile Glu Tyr Arg Leu Ser Val Gly Thr Asn Glu Lys Ile

305 310 315 320
 Gln Leu Ser Ser Leu Ile Ala Ala Phe Gln Val Thr Arg Asp Leu Ile
 325 330 335
 Val Ala Glu Ala
 340

(2) ANGABEN ZU SEQ ID NO: 2:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 319 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: Archaeoglobus fulgidus

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 2:

Met Glu Asn Phe Glu Ile Trp Val Glu Lys Tyr Arg Pro Arg Thr Leu
 1 5 10 15
 Asp Glu Val Val Gly Gln Asp Glu Val Ile Gln Arg Leu Lys Gly Tyr
 20 25 30
 Val Glu Arg Lys Asn Ile Pro His Leu Leu Phe Ser Gly Pro Pro Gly
 35 40 45
 Thr Gly Lys Thr Ala Thr Ala Ile Ala Leu Ala Arg Asp Leu Phe Gly
 50 55 60
 Glu Asn Trp Arg Asp Asn Phe Ile Glu Met Asn Ala Ser Asp Glu Arg
 65 70 75 80
 Gly Ile Asp Val Val Arg His Lys Ile Lys Glu Phe Ala Arg Thr Ala
 85 90 95
 Pro Ile Gly Gly Ala Pro Phe Lys Ile Ile Phe Leu Asp Glu Ala Asp
 100 105 110
 Ala Leu Thr Ala Asp Ala Gln Ala Ala Leu Arg Arg Thr Met Glu Met
 115 120 125

Tyr Ser Lys Ser Cys Arg Phe Ile Leu Ser Cys Asn Tyr Val Ser Arg
 130 135 140
 Ile Ile Glu Pro Ile Gln Ser Arg Cys Ala Val Phe Arg Phe Lys Pro
 145 150 155 160
 Val Pro Lys Glu Ala Met Lys Lys Arg Leu Leu Glu Ile Cys Glu Lys
 165 170 175
 Glu Gly Val Lys Ile Thr Glu Asp Gly Leu Glu Ala Leu Ile Tyr Ile
 180 185 190
 Ser Gly Gly Asp Phe Arg Lys Ala Ile Asn Ala Leu Gln Gly Ala Ala
 195 200 205
 Ala Ile Gly Glu Val Val Asp Ala Asp Thr Ile Tyr Gln Ile Thr Ala
 210 215 220
 Thr Ala Arg Pro Glu Glu Met Thr Glu Leu Ile Gln Thr Ala Leu Lys
 225 230 235 240
 Gly Asn Phe Met Glu Ala Arg Glu Leu Leu Asp Arg Leu Met Val Glu
 245 250 255
 Tyr Gly Met Ser Gly Glu Asp Ile Val Ala Gln Leu Phe Arg Glu Ile
 260 265 270
 Ile Ser Met Pro Ile Lys Asp Ser Leu Lys Val Gln Leu Ile Asp Lys
 275 280 285
 Leu Gly Glu Val Asp Phe Arg Leu Thr Glu Gly Ala Asn Glu Arg Ile
 290 295 300
 Gln Leu Asp Ala Tyr Leu Ala Tyr Leu Ser Thr Leu Ala Lys Lys
 305 310 315

(2) ANGABEN ZU SEQ ID NO: 3:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 1847 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: *Methanococcus jannaschii*

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 3:

Met Val Ile Ile Met Glu Lys Pro Trp Val Glu Lys Tyr Arg Pro Lys
 1 5 10 15

Thr Leu Asp Asp Ile Val Gly Gln Asp Glu Ile Val Lys Arg Leu Lys
 20 25 30

Lys Tyr Val Glu Lys Lys Ser Met Pro His Leu Leu Phe Ser Gly Pro
 35 40 45

Pro Gly Val Gly Lys Cys Leu Thr Gly Asp Thr Lys Val Ile Val Asn
 50 55 60

Gly Glu Ile Arg Glu Ile Gly Glu Val Ile Glu Glu Ile Ser Asn Gly
 65 70 75 80

Lys Phe Gly Val Thr Leu Thr Asn Asn Leu Lys Val Leu Gly Ile Asp
 85 90 95

Glu Asp Gly Lys Ile Arg Glu Phe Asp Val Gln Tyr Val Tyr Lys Asp
 100 105 110

Lys Thr Asn Thr Leu Ile Lys Ile Lys Thr Lys Met Gly Arg Glu Leu
 115 120 125

Lys Val Thr Thr Tyr His Pro Leu Leu Ile Asn His Lys Asn Gly Glu
 130 135 140

Ile Lys Trp Glu Lys Ala Glu Asn Leu Lys Val Gly Asp Lys Leu Ala
 145 150 155 160

Thr Pro Arg Tyr Ile Leu Phe Asn Glu Ser Asp Tyr Asn Glu Glu Leu
 165 170 175

Ala Glu Trp Leu Gly Tyr Phe Ile Gly Asp Gly His Ala Asp Lys Glu
 180 185 190

Ser Asn Lys Ile Thr Phe Thr Asn Gly Asp Glu Lys Leu Arg Lys Arg
 195 200 205

Phe Ala Glu Leu Thr Glu Lys Leu Phe Lys Asp Ala Lys Ile Lys Glu
 210 215 220

Arg Ile His Lys Asp Arg Thr Pro Asp Ile Tyr Val Asn Ser Lys Glu
 225 230 235 240

Ala Val Glu Phe Ile Asp Lys Leu Gly Leu Arg Gly Lys Lys Ala Asp

	245	250	255
Lys Val Arg Ile Pro Lys Glu Ile Met Arg Ser Asp Ala Leu Arg Ala			
	260	265	270
Phe Leu Arg Ala Tyr Phe Asp Cys Asp Gly Gly Ile Glu Lys His Ser			
	275	280	285
Ile Val Leu Ser Thr Ala Ser Lys Glu Met Ala Glu Asp Leu Val Tyr			
	290	295	300
Ala Leu Leu Arg Phe Gly Ile Ile Ala Lys Leu Arg Glu Lys Val Asn			
	305	310	315
Lys Asn Asn Asn Lys Val Tyr Tyr His Ile Val Ile Ser Asn Ser Ser			
	325	330	335
Asn Leu Arg Thr Phe Leu Asp Asn Ile Gly Phe Ser Gln Glu Arg Lys			
	340	345	350
Leu Lys Lys Leu Leu Glu Ile Ile Lys Asp Glu Asn Pro Asn Leu Asp			
	355	360	365
Val Ile Thr Ile Asp Lys Glu Lys Ile Arg Tyr Ile Arg Asp Arg Leu			
	370	375	380
Lys Val Lys Leu Thr Arg Asp Ile Glu Lys Asp Asn Trp Ser Tyr Asn			
	385	390	395
Lys Cys Arg Lys Ile Thr Gln Glu Leu Leu Lys Glu Ile Tyr Tyr Arg			
	405	410	415
Leu Glu Glu Leu Lys Glu Ile Glu Lys Ala Leu Glu Glu Asn Ile Leu			
	420	425	430
Ile Asp Trp Asp Glu Val Ala Glu Arg Arg Lys Glu Ile Ala Glu Lys			
	435	440	445
Thr Gly Ile Arg Ser Asp Arg Ile Leu Glu Tyr Ile Arg Gly Lys Arg			
	450	455	460
Lys Pro Ser Leu Lys Asn Tyr Ile Lys Ile Ala Asn Thr Leu Gly Lys			
	465	470	475
Asn Ile Glu Lys Ile Ile Asp Ala Met Arg Ile Phe Ala Lys Lys Tyr			
	485	490	495
Ser Ser Tyr Ala Glu Ile Gly Lys Met Leu Asn Met Trp Asn Ser Ser			
	500	505	510

Ile Lys Ile Tyr Leu Glu Ser Asn Thr Gln Glu Ile Glu Lys Leu Glu
 515 520 525

Glu Ile Arg Lys Thr Glu Leu Lys Leu Val Lys Glu Ile Leu Asn Asp
 530 535 540

Glu Lys Leu Ile Asp Ser Ile Gly Tyr Val Leu Phe Leu Ala Ser Asn
 545 550 555 560

Glu Ile Tyr Trp Asp Glu Ile Val Glu Ile Glu Gln Leu Asn Gly Glu
 565 570 575

Phe Thr Ile Tyr Asp Leu His Val Pro Arg Tyr His Asn Phe Ile Gly
 580 585 590

Gly Asn Leu Pro Thr Ile Leu His Asn Thr Thr Ala Ala Leu Cys Leu
 595 600 605

Ala Arg Asp Leu Phe Gly Glu Asn Trp Arg Asp Asn Phe Leu Glu Leu
 610 615 620

Asn Ala Ser Val Ser Lys Asp Thr Pro Ile Leu Val Lys Ile Asp Gly
 625 630 635 640

Lys Val Lys Arg Thr Thr Phe Glu Glu Leu Asp Lys Ile Tyr Phe Glu
 645 650 655

Thr Asn Asp Glu Asn Glu Met Tyr Lys Lys Val Asp Asn Leu Glu Val
 660 665 670

Leu Thr Val Asp Glu Asn Phe Arg Val Arg Trp Arg Lys Val Ser Thr
 675 680 685

Ile Ile Arg His Lys Val Asp Lys Ile Leu Arg Ile Lys Phe Glu Gly
 690 695 700

Gly Tyr Ile Glu Leu Thr Gly Asn His Ser Ile Met Met Leu Asp Glu
 705 710 715 720

Asn Gly Leu Val Ala Lys Lys Ala Ser Asp Ile Lys Val Gly Asp Cys
 725 730 735

Phe Leu Ser Phe Val Ala Asn Ile Glu Gly Glu Lys Asp Arg Leu Asp
 740 745 750

Leu Lys Glu Phe Glu Pro Lys Asp Ile Thr Ser Arg Val Lys Ile Ile
 755 760 765

Asn Asp Phe Asp Ile Asp Glu Asp Thr Ala Trp Met Leu Gly Leu Tyr

770	775	780	
Val Ala Glu Gly	Ala Val Gly Phe	Lys Gly Lys Thr	Ser Gly Gln Val
785	790	795	800
Ile Tyr Thr Leu	Gly Ser His Glu	His Asp Leu	Ile Asn Lys Leu Asn
	805	810	815
Asp Ile Val Asp	Lys Lys Gly Phe	Ser Lys Tyr Glu	Asn Phe Thr Gly
	820	825	830
Ser Gly Phe Asp	Arg Lys Arg Leu	Ser Ala Lys Gln	Ile Arg Ile Leu
	835	840	845
Asn Thr Gln Leu	Ala Arg Phe Val	Glu Glu Asn Phe	Tyr Asp Gly Asn
	850	855	860
Gly Arg Arg Ala	Arg Asn Lys Arg	Ile Pro Asp Ile	Ile Phe Glu Leu
	865	870	875
Lys Glu Asn Leu	Arg Val Glu Phe	Leu Lys Gly Leu	Ala Asp Gly Asp
	885	890	895
Ser Ser Gly Asn	Trp Arg Glu Val	Val Arg Ile Ser	Ser Lys Ser Asp
	900	905	910
Asn Leu Leu Ile	Asp Thr Val Trp	Leu Ala Arg Ile	Ser Gly Ile Glu
	915	920	925
Ser Ser Ile Phe	Glu Asn Glu Ala	Arg Leu Ile Trp	Lys Gly Gly Met
	930	935	940
Lys Trp Lys Lys	Ser Asn Leu Leu	Pro Ala Glu Pro	Ile Ile Lys Met
	945	950	955
Ile Lys Lys Leu	Glu Asn Lys Ile	Asn Gly Asn Trp	Arg Tyr Ile Leu
	965	970	975
Arg His Gln Leu	Tyr Glu Gly Lys	Lys Arg Val Ser	Lys Asp Lys Ile
	980	985	990
Lys Gln Ile Leu	Glu Met Val Asn	Val Glu Lys Leu	Ser Asp Lys Glu
	995	1000	1005
Lys Glu Val Tyr	Asp Leu Leu Lys	Lys Leu Ser Lys	Thr Glu Leu Tyr
	1010	1015	1020
Ala Leu Val Val	Lys Glu Ile Glu	Ile Ile Asp Tyr	Asn Asp Phe Val
	1025	1030	1035
			1040

Tyr Asp Val Ser Val Pro Asn Asn Glu Met Phe Phe Ala Gly Asn Val
 1045 1050 1055
 Pro Ile Leu Leu His Asn Ser Asp Glu Arg Gly Ile Asp Val ile Arg
 1060 1065 1070
 Thr Lys Val Lys Asp Phe Ala Arg Thr Lys Pro Ile Gly Asp Val Pro
 1075 1080 1085
 Phe Lys Ile Ile Phe Leu Asp Glu Ser Asp Ala Leu Thr Ala Asp Ala
 1090 1095 1100
 Gln Asn Ala Leu Arg Arg Thr Met Glu Lys Tyr Ser Asp Val Cys Arg
 1105 1110 1115 1120
 Phe Ile Leu Ser Cys Leu Thr Gly Asp Ala Lys Ile Thr Leu Pro Asp
 1125 1130 1135
 Glu Arg Glu Ile Lys Ile Glu Asp Phe Ile Lys Met Phe Glu Glu Arg
 1140 1145 1150
 Lys Leu Lys His Val Leu Asn Arg Asn Gly Glu Asp Leu Val Leu Ala
 1155 1160 1165
 Gly Val Lys Phe Asn Ser Lys Ile Val Asn His Lys Val Tyr Arg Leu
 1170 1175 1180
 Val Leu Glu Ser Gly Arg Glu Ile Glu Ala Thr Gly Asp His Lys Phe
 1185 1190 1195 1200
 Leu Thr Arg Asp Gly Trp Lys Glu Val Tyr Glu Leu Lys Glu Asp Asp
 1205 1210 1215
 Glu Val Leu Val Tyr Pro Ala Leu Glu Gly Val Gly Phe Glu Val Asp
 1220 1225 1230
 Glu Arg Arg Ile Ile Gly Leu Asn Glu Phe Tyr Glu Phe Leu Thr Asn
 1235 1240 1245
 Tyr Glu Ile Lys Leu Gly Tyr Lys Pro Leu Gly Lys Ala Lys Ser Tyr
 1250 1255 1260
 Lys Glu Leu Ile Thr Arg Asp Lys Glu Lys Ile Leu Ser Arg Val Leu
 1265 1270 1275 1280
 Glu Leu Ser Asp Lys Tyr Ser Lys Ser Glu Ile Arg Arg Lys Ile Glu
 1285 1290 1295
 Glu Glu Phe Gly Ile Lys Ile Ser Leu Thr Thr Ile Lys Asn Leu Ile

1300	1305	1310
Asn Gly Lys Ile Asp Gly Phe Ala Leu Lys Tyr Val Arg Lys Ile Lys		
1315	1320	1325
Glu Leu Gly Trp Asp Glu Ile Thr Tyr Asp Asp Glu Lys Ala Gly Ile		
1330	1335	1340
Phe Ala Arg Leu Leu Gly Phe Ile Ile Gly Asp Gly His Leu Ser Lys		
1345	1350	1355
Ser Lys Glu Gly Arg Ile Leu Ile Thr Ala Thr Ile Asn Glu Leu Glu		
1365	1370	1375
Gly Ile Lys Lys Asp Leu Glu Lys Leu Gly Ile Lys Ala Ser Asn Ile		
1380	1385	1390
Ile Glu Lys Asp Ile Glu His Lys Leu Asp Gly Arg Glu Ile Lys Gly		
1395	1400	1405
Lys Thr Ser Phe Ile Tyr Ile Asn Asn Lys Ala Phe Tyr Leu Leu Leu		
1410	1415	1420
Asn Phe Trp Gly Val Glu Ile Gly Asn Lys Thr Ile Asn Gly Tyr Asn		
1425	1430	1435
Ile Pro Lys Trp Ile Lys Tyr Gly Asn Lys Phe Val Lys Arg Glu Phe		
1445	1450	1455
Leu Arg Gly Leu Phe Gly Ala Asp Gly Thr Lys Pro Tyr Ile Lys Lys		
1460	1465	1470
Tyr Asn Ile Asn Gly Ile Lys Leu Gly Ile Arg Val Glu Asn Ile Ser		
1475	1480	1485
Lys Asp Lys Thr Leu Glu Phe Phe Glu Glu Val Lys Lys Met Leu Glu		
1490	1495	1500
Glu Phe Glu Val Glu Ser Tyr Ile Lys Val Ser Lys Ile Asp Asn Lys		
1505	1510	1515
Asn Leu Thr Glu Leu Ile Val Lys Ala Asn Asn Lys Asn Tyr Leu Lys		
1525	1530	1535
Tyr Leu Ser Arg Ile Ser Tyr Ala Tyr Glu Lys Asp Asn Phe Ala Arg		
1540	1545	1550
Leu Val Gly Glu Tyr Leu Arg Ile Lys Glu Ala Tyr Lys Asp Ile Ile		
1555	1560	1565

Leu Lys Glu Ile Ala Glu Asn Ala Leu Lys Glu Ala Asp Gly Glu Lys
 1570 1575 1580

Ser Leu Arg Glu Leu Ala Arg Lys Tyr Asn Val Pro Val Asp Phe Ile
 1585 1590 1595 1600

Ile Asn Gln Leu Lys Gly Lys Asp Ile Gly Leu Pro Arg Asn Phe Met
 1605 1610 1615

Thr Phe Glu Glu Phe Leu Lys Glu Lys Val Val Asp Gly Lys Tyr Val
 1620 1625 1630

Ser Glu Arg Ile Ile Lys Lys Glu Cys Ile Gly Tyr Arg Asp Val Tyr
 1635 1640 1645

Asp Ile Thr Cys His Lys Asp Pro Ser Phe Ile Ala Asn Gly Phe Val
 1650 1655 1660

Ser His Asn Cys Asn Tyr Pro Ser Lys Ile Ile Pro Pro Ile Gln Ser
 1665 1670 1675 1680

Arg Cys Ala Val Phe Arg Phe Ser Pro Leu Lys Lys Glu Asp Ile Ala
 1685 1690 1695

Lys Lys Leu Lys Glu Ile Ala Glu Lys Glu Gly Leu Asn Leu Thr Glu
 1700 1705 1710

Ser Gly Leu Glu Ala Ile Ile Tyr Val Ser Glu Gly Asp Met Arg Lys
 1715 1720 1725

Ala Ile Asn Val Leu Gln Thr Ala Ala Ala Leu Ser Asp Val Ile Asp
 1730 1735 1740

Asp Glu Ile Val Tyr Lys Val Ser Ser Arg Ala Arg Pro Glu Glu Val
 1745 1750 1755 1760

Lys Lys Met Met Glu Leu Ala Leu Asp Gly Lys Phe Met Glu Ala Arg
 1765 1770 1775

Asp Leu Leu Tyr Lys Leu Met Val Glu Trp Gly Met Ser Gly Glu Asp
 1780 1785 1790

Ile Leu Asn Gln Met Phe Arg Glu Ile Asn Ser Leu Asp Ile Asp Glu
 1795 1800 1805

Arg Lys Lys Val Glu Leu Ala Asp Ala Ile Gly Glu Thr Asp Phe Arg
 1810 1815 1820

Ile Val Glu Gly Ala Asn Glu Arg Ile Gln Leu Ser Ala Leu Leu Ala

1825 1830 1835 1840

Lys Met Ala Leu Met Gly Arg
1845

(2) ANGABEN ZU SEQ ID NO: 4:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 855 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: *Pyrococcus horikoshii*

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 4:

Met His Asn Met Glu Glu Val Arg Glu Val Lys Val Leu Glu Lys Pro
1 5 10 15

Trp Val Glu Lys Tyr Arg Pro Gln Arg Leu Asp Glu Ile Val Gly Gln
20 25 30

Glu His Ile Val Lys Arg Leu Lys His Tyr Val Lys Thr Gly Ser Met
35 40 45

Pro His Leu Leu Phe Ala Gly Pro Pro Gly Val Gly Lys Cys Leu Thr
50 55 60

Gly Asp Thr Lys Val Ile Ala Asn Gly Gln Leu Phe Glu Leu Arg Glu
65 70 75 80

Leu Val Glu Lys Ile Ser Gly Gly Lys Phe Gly Pro Thr Pro Val Lys
85 90 95

Gly Leu Lys Val Ile Gly Ile Asp Glu Asp Gly Lys Leu Arg Glu Phe
100 105 110

Glu Val Gln Tyr Val Tyr Lys Asp Lys Thr Glu Arg Leu Ile Arg Ile
115 120 125

Arg Thr Arg Leu Gly Arg Glu Leu Lys Val Thr Pro Tyr His Pro Leu
130 135 140

Leu Val Asn Arg Arg Asn Gly Glu Ile Lys Trp Val Lys Ala Glu Glu
145 150 155 160

Leu Lys Pro Gly Asp Lys Leu Ala Val Pro Arg Phe Leu Pro Ile Val
165 170 175

Thr Gly Glu Asp Pro Leu Ala Glu Trp Leu Gly Tyr Phe Leu Gly Gly
180 185 190

Gly Tyr Ala Asp Ser Lys Glu Asn Leu Ile Met Phe Thr Asn Glu Asp
195 200 205

Pro Leu Leu Arg Gln Arg Phe Met Glu Leu Thr Glu Lys Leu Phe Ser
210 215 220

Asp Ala Arg Ile Arg Glu Ile Thr His Glu Asn Gly Thr Ser Lys Val
225 230 235 240

Tyr Val Asn Ser Lys Lys Ala Leu Lys Leu Val Asn Ser Leu Gly Asn
245 250 255

Ala His Ile Pro Lys Glu Cys Trp Arg Gly Ile Arg Ser Phe Leu Arg
260 265 270

Ala Tyr Phe Asp Cys Asn Gly Gly Val Lys Gly Asn Ala Ile Val Leu
275 280 285

Ala Thr Ala Ser Lys Glu Met Ser Gln Glu Ile Ala Tyr Ala Leu Ala
290 295 300

Gly Phe Gly Ile Ile Ser Arg Ile Gln Glu Tyr Arg Val Ile Ile Ser
305 310 315 320

Gly Ser Asp Asn Val Lys Lys Phe Leu Asn Glu Ile Gly Phe Ile Asn
325 330 335

Arg Asn Lys Leu Glu Lys Ala Leu Lys Leu Val Lys Lys Asp Asp Pro
340 345 350

Gly His Asp Gly Leu Glu Ile Asn Tyr Glu Leu Ile Ser Tyr Val Lys
355 360 365

Asp Arg Leu Arg Leu Ser Phe Phe Asn Asp Lys Arg Ser Trp Ser Tyr
370 375 380

Arg Glu Ala Lys Glu Ile Ser Trp Glu Leu Met Lys Glu Ile Tyr Tyr
385 390 395 400

Arg Leu Asp Glu Leu Glu Lys Leu Lys Glu Ser Leu Ser Arg Gly Ile

	405	410	415
Leu Ile Asp Trp Asn Glu Val Ala Lys Arg Ile Glu Glu Val Ala Glu			
	420	425	430
Glu Thr Gly Ile Arg Ala Asp Glu Leu Leu Glu Tyr Ile Glu Gly Lys			
	435	440	445
Arg Lys Leu Ser Phe Lys Asp Tyr Ile Lys Ile Ala Lys Val Leu Gly			
	450	455	460
Ile Asp Val Glu His Thr Ile Glu Ala Met Arg Val Phe Ala Arg Lys			
	465	470	475
			480
Tyr Ser Ser Tyr Ala Glu Ile Gly Arg Arg Leu Gly Thr Trp Asn Ser			
	485	490	495
Ser Val Lys Thr Ile Leu Glu Ser Asn Ala Val Asn Val Glu Ile Leu			
	500	505	510
Glu Arg Ile Arg Lys Ile Glu Leu Glu Leu Ile Glu Glu Ile Leu Ser			
	515	520	525
Asp Glu Lys Leu Lys Glu Gly Ile Ala Tyr Leu Ile Phe Leu Ser Gln			
	530	535	540
Asn Glu Leu Tyr Trp Asp Glu Ile Thr Lys Val Glu Glu Leu Arg Gly			
	545	550	555
			560
Glu Phe Ile Ile Tyr Asp Leu His Val Pro Gly Tyr His Asn Phe Ile			
	565	570	575
Ala Gly Asn Met Pro Thr Val Val His Asn Thr Thr Ala Ala Leu Ala			
	580	585	590
Leu Ser Arg Glu Leu Phe Gly Glu Asn Trp Arg His Asn Phe Leu Glu			
	595	600	605
Leu Asn Ala Ser Asp Glu Arg Gly Ile Asn Val Ile Arg Glu Lys Val			
	610	615	620
Lys Glu Phe Ala Arg Thr Lys Pro Ile Gly Gly Ala Ser Phe Lys Ile			
	625	630	635
			640
Ile Phe Leu Asp Glu Ala Asp Ala Leu Thr Gln Asp Ala Gln Gln Ala			
	645	650	655
Leu Arg Arg Thr Met Glu Met Phe Ser Ser Asn Val Arg Phe Ile Leu			
	660	665	670

Ser Cys Asn Tyr Ser Ser Lys Ile Ile Glu Pro Ile Gln Ser Arg Cys
 675 680 685
 Ala Ile Phe Arg Phe Arg Pro Leu Arg Asp Glu Asp Ile Ala Lys Arg
 690 695 700
 Leu Arg Tyr Ile Ala Glu Asn Glu Gly Leu Glu Leu Thr Glu Glu Gly
 705 710 715 720
 Leu Gln Ala Ile Leu Tyr Ile Ala Glu Gly Asp Met Arg Arg Ala Ile
 725 730 735
 Asn Ile Leu Gln Ala Ala Ala Ala Leu Asp Lys Lys Ile Thr Asp Glu
 740 745 750
 Asn Val Phe Met Val Ala Ser Arg Ala Arg Pro Glu Asp Ile Arg Glu
 755 760 765
 Met Met Leu Leu Ala Leu Lys Gly Asn Phe Leu Lys Ala Arg Glu Lys
 770 775 780
 Leu Arg Glu Ile Leu Leu Lys Gln Gly Leu Ser Gly Glu Asp Val Leu
 785 790 795 800
 Ile Gln Met His Lys Glu Val Phe Asn Leu Pro Ile Asp Glu Pro Thr
 805 810 815
 Lys Val Tyr Leu Ala Asp Lys Ile Gly Glu Tyr Asn Phe Arg Leu Val
 820 825 830
 Glu Gly Ala Asn Glu Met Ile Gln Leu Glu Ala Leu Leu Ala Gln Phe
 835 840 845
 Thr Leu Val Gly Lys Lys Lys
 850 855

(2) ANGABEN ZU SEQ ID NO: 5:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 321 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: Methanobacterium thermoautotrophicum

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 5:

Met Ile Ile Met Asn Gly Pro Trp Val Glu Lys Tyr Arg Pro Gln Lys
 1 5 10 15
 Leu Asp Asp Ile Val Gly Gln Glu His Ile Ile Pro Arg Leu Lys Arg
 20 25 30
 Tyr Val Glu Glu Lys Ser Met Pro Asn Leu Met Phe Thr Gly Pro Ala
 35 40 45
 Gly Val Gly Lys Thr Thr Ala Ala Leu Ala Leu Ala Arg Glu Ile Leu
 50 55 60
 Gly Glu Tyr Trp Arg Gln Asn Phe Leu Glu Leu Asn Ala Ser Asp Ala
 65 70 75 80
 Arg Gly Ile Asp Thr Val Arg Thr Ser Ile Lys Asn Phe Cys Arg Leu
 85 90 95
 Lys Pro Val Gly Ala Pro Phe Arg Ile Ile Phe Leu Asp Glu Val Asp
 100 105 110
 Asn Met Thr Lys Asp Ala Gln His Ala Leu Arg Arg Glu Met Glu Met
 115 120 125
 Tyr Thr Lys Thr Ser Ser Phe Ile Leu Ser Cys Asn Tyr Ser Ser Lys
 130 135 140
 Ile Ile Asp Pro Ile Gln Ser Arg Cys Ala Ile Phe Arg Phe Leu Pro
 145 150 155 160
 Leu Lys Gly His Gln Ile Ile Lys Arg Leu Glu Tyr Ile Ala Glu Lys
 165 170 175
 Glu Asn Leu Glu Tyr Glu Ala His Ala Leu Glu Thr Ile Val Tyr Phe
 180 185 190
 Ala Glu Gly Asp Leu Arg Lys Ala Ile Asn Leu Leu Gln Ser Ala Ala
 195 200 205
 Ser Leu Gly Glu Lys Ile Thr Glu Ser Ser Ile Tyr Asp Val Val Ser
 210 215 220
 Arg Ala Arg Pro Lys Asp Val Arg Lys Met Ile Lys Thr Ile Leu Asp
 225 230 235 240
 Gly Lys Phe Met Glu Ala Arg Asp Met Leu Arg Glu Ile Met Val Leu

245 250 255
 Gln Gly Ile Ser Gly Glu Asp Met Val Thr Gln Ile Tyr Gln Glu Leu
 260 265 270
 Ser Arg Leu Ala Met Glu Gly Glu Val Asp Gly Asp Arg Tyr Val Gly
 275 280 285
 Leu Ile Asp Ala Ile Gly Glu Tyr Asp Phe Arg Ile Arg Glu Gly Ala
 290 295 300
 Asn Pro Arg Ile Gln Leu Glu Ala Leu Leu Ala Arg Phe Leu Glu His
 305 310 315 320
 Ala

(2) ANGABEN ZU SEQ ID NO: 6:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 1148 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: Homo sapiens

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 6:

Met Asp Ile Arg Lys Phe Phe Gly Val Ile Pro Ser Gly Lys Lys Leu
 1 5 10 15
 Val Ser Glu Thr Val Lys Lys Asn Glu Lys Thr Lys Ser Asp Glu Glu
 20 25 30
 Thr Leu Lys Ala Lys Lys Gly Ile Lys Glu Ile Lys Val Asn Ser Ser
 35 40 45
 Arg Lys Glu Asp Asp Phe Lys Gln Lys Gln Pro Ser Lys Lys Lys Arg
 50 55 60
 Ile Ile Tyr Asp Ser Asp Ser Glu Ser Glu Glu Thr Leu Gln Val Lys
 65 70 75 80

Asn Ala Lys Lys Pro Pro Glu Lys Leu Pro Val Ser Ser Lys Pro Gly
 85 90 95

Lys Ile Ser Arg Gln Asp Pro Val Thr Tyr Ile Ser Glu Thr Asp Glu
 100 105 110

Glu Asp Asp Phe Met Cys Lys Lys Ala Ala Ser Lys Ser Lys Glu Asn
 115 120 125

Gly Arg Ser Thr Asn Ser His Leu Gly Thr Ser Asn Met Lys Lys Asn
 130 135 140

Glu Glu Asn Thr Lys Thr Lys Asn Lys Pro Leu Ser Pro Ile Lys Leu
 145 150 155 160

Thr Pro Thr Ser Val Leu Asp Tyr Phe Gly Thr Gly Ser Val Gln Arg
 165 170 175

Ser Asn Lys Lys Met Val Ala Ser Lys Arg Lys Glu Leu Ser Gln Asn
 180 185 190

Thr Asp Glu Ser Gly Leu Asn Asp Glu Ala Ile Ala Lys Gln Leu Gln
 195 200 205

Leu Asp Glu Asp Ala Glu Leu Glu Arg Gln Leu His Glu Asp Glu Glu
 210 215 220

Phe Ala Arg Thr Leu Ala Met Leu Asp Glu Glu Pro Lys Thr Lys Lys
 225 230 235 240

Ala Arg Lys Asp Thr Glu Ala Gly Glu Thr Phe Ser Ser Val Gln Ala
 245 250 255

Asn Leu Ser Lys Ala Glu Lys His Lys Tyr Pro His Lys Val Lys Thr
 260 265 270

Ala Gln Val Ser Asp Glu Arg Lys Ser Tyr Ser Pro Arg Lys Gln Ser
 275 280 285

Lys Tyr Glu Ser Ser Lys Glu Ser Gln Gln His Ser Lys Ser Ser Ala
 290 295 300

Asp Lys Ile Gly Glu Val Ser Ser Pro Lys Ala Ser Ser Lys Leu Ala
 305 310 315 320

Ile Met Lys Arg Lys Lys Glu Ser Ser Tyr Lys Glu Ile Glu Pro Val
 325 330 335

Ala Ser Lys Arg Lys Glu Asn Ala Ile Lys Leu Lys Gly Glu Thr Lys

340 345 350
 Thr Pro Lys Lys Thr Lys Ser Ser Pro Ala Lys Lys Glu Ser Val Ser
 355 360 365
 Pro Glu Asp Ser Glu Lys Lys Arg Thr Asn Tyr Gln Ala Tyr Arg Ser
 370 375 380
 Tyr Leu Asn Arg Glu Gly Pro Lys Ala Leu Gly Ser Lys Glu Ile Pro
 385 390 395 400
 Lys Gly Ala Glu Asn Cys Leu Glu Gly Leu Ile Phe Val Ile Thr Gly
 405 410 415
 Val Leu Glu Ser Ile Glu Arg Asp Glu Ala Lys Ser Leu Ile Glu Arg
 420 425 430
 Tyr Gly Gly Lys Val Thr Gly Asn Val Ser Lys Lys Thr Asn Tyr Leu
 435 440 445
 Val Met Gly Arg Asp Ser Gly Gln Ser Lys Ser Asp Lys Ala Ala Ala
 450 455 460
 Leu Gly Thr Lys Ile Ile Asp Glu Asp Gly Leu Leu Asn Leu Ile Arg
 465 470 475 480
 Thr Met Pro Gly Lys Lys Ser Lys Tyr Glu Ile Ala Val Glu Thr Glu
 485 490 495
 Met Lys Lys Glu Ser Lys Leu Glu Arg Thr Pro Gln Lys Asn Val Gln
 500 505 510
 Gly Lys Arg Lys Ile Ser Pro Ser Lys Lys Glu Ser Glu Ser Lys Lys
 515 520 525
 Ser Arg Pro Thr Ser Lys Arg Asp Ser Leu Ala Lys Thr Ile Lys Lys
 530 535 540
 Glu Thr Asp Val Phe Trp Lys Ser Leu Asp Phe Lys Glu Gln Val Ala
 545 550 555 560
 Glu Glu Thr Ser Gly Asp Ser Lys Ala Arg Asn Leu Ala Asp Asp Ser
 565 570 575
 Ser Glu Asn Lys Val Glu Asn Leu Leu Trp Val Asp Lys Tyr Lys Pro
 580 585 590
 Thr Ser Leu Lys Thr Ile Ile Gly Gln Gln Gly Asp Gln Ser Cys Ala
 595 600 605

Asn Lys Leu Leu Arg Trp Leu Arg Asn Trp Gln Lys Ser Ser Ser Glu
610 615 620

Asp Lys Lys His Ala Ala Lys Phe Gly Lys Phe Ser Gly Lys Asp Asp
625 630 635 640

Gly Ser Ser Phe Lys Ala Ala Leu Leu Ser Gly Pro Pro Gly Val Gly
645 650 655

Lys Thr Thr Thr Ala Ser Leu Val Cys Gln Glu Leu Gly Tyr Ser Tyr
660 665 670

Val Glu Leu Asn Ala Ser Asp Thr Arg Ser Lys Ser Ser Leu Lys Ala
675 680 685

Ile Val Ala Glu Ser Leu Asn Asn Thr Ser Ile Lys Gly Phe Tyr Ser
690 695 700

Asn Gly Ala Ala Ser Ser Val Ser Thr Lys His Ala Leu Ile Met Asp
705 710 715 720

Glu Val Asp Gly Met Ala Gly Asn Glu Asp Arg Gly Gly Ile Gln Glu
725 730 735

Leu Ile Gly Leu Ile Lys His Thr Lys Ile Pro Ile Ile Cys Met Cys
740 745 750

Asn Asp Arg Asn His Pro Lys Ile Arg Ser Leu Val His Tyr Cys Phe
755 760 765

Asp Leu Arg Phe Gln Arg Pro Arg Val Glu Gln Ile Lys Gly Ala Met
770 775 780

Met Ser Ile Ala Phe Lys Glu Gly Leu Lys Ile Pro Pro Pro Ala Met
785 790 795 800

Asn Glu Ile Ile Leu Gly Ala Asn Gln Asp Ile Arg Gln Val Leu His
805 810 815

Asn Leu Ser Met Trp Cys Ala Arg Ser Lys Ala Leu Thr Tyr Asp Gln
820 825 830

Ala Lys Ala Asp Ser His Arg Ala Lys Lys Asp Ile Lys Met Gly Pro
835 840 845

Phe Asp Val Ala Arg Lys Val Phe Ala Ala Gly Glu Glu Thr Ala His
850 855 860

Met Ser Leu Val Asp Lys Ser Asp Leu Phe Phe His Asp Tyr Ser Ile

865 870 875 880
 Ala Pro Leu Phe Val Gln Glu Asn Tyr Ile His Val Lys Pro Val Ala
 885 890 895
 Ala Gly Gly Asp Met Lys Lys His Leu Met Leu Leu Ser Arg Ala Ala
 900 905 910
 Asp Ser Ile Cys Asp Gly Asp Leu Val Asp Ser Gln Ile Arg Ser Lys
 915 920 925
 Gln Asn Trp Ser Leu Leu Pro Ala Gln Ala Ile Tyr Ala Ser Val Leu
 930 935 940
 Pro Gly Glu Leu Met Arg Gly Tyr Met Thr Gln Phe Pro Thr Phe Pro
 945 950 955 960
 Ser Trp Leu Gly Lys His Ser Ser Thr Gly Lys His Asp Arg Ile Val
 965 970 975
 Gln Asp Leu Ala Leu His Met Ser Leu Arg Thr Tyr Ser Ser Lys Arg
 980 985 990
 Thr Val Asn Met Asp Tyr Leu Ser Leu Leu Arg Asp Ala Leu Val Gln
 995 1000 1005
 Pro Leu Thr Ser Gln Gly Val Asp Gly Val Gln Asp Val Val Ala Leu
 1010 1015 1020
 Met Asp Thr Tyr Tyr Leu Met Lys Glu Asp Phe Glu Asn Ile Met Glu
 1025 1030 1035 1040
 Ile Ser Ser Trp Gly Gly Lys Pro Ser Pro Phe Ser Lys Leu Asp Pro
 1045 1050 1055
 Lys Val Lys Ala Ala Phe Thr Arg Ala Tyr Asn Lys Glu Ala His Leu
 1060 1065 1070
 Thr Pro Tyr Ser Leu Gln Ala Ile Lys Ala Ser Arg His Ser Thr Ser
 1075 1080 1085
 Pro Ser Leu Asp Ser Glu Tyr Asn Glu Glu Leu Asn Glu Asp Asp Ser
 1090 1095 1100
 Gln Ser Asp Glu Lys Asp Gln Asp Ala Ile Glu Thr Asp Ala Met Ile
 1105 1110 1115 1120
 Lys Lys Lys Thr Lys Ser Ser Lys Pro Ser Lys Pro Glu Lys Asp Lys
 1125 1130 1135

Glu Pro Arg Lys Gly Lys Gly Lys Ser Ser Lys Lys
 1140 1145

(2) ANGABEN ZU SEQ ID NO: 7:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 479 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: Archaeoglobus fulgidus

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 7:

Met Leu Trp Val Glu Lys Tyr Arg Pro Lys Thr Leu Glu Glu Val Val
 1 5 10 15

Ala Asp Lys Ser Ile Ile Thr Arg Val Ile Lys Trp Ala Lys Ser Trp
 20 25 30

Lys Arg Gly Ser Lys Pro Leu Leu Leu Ala Gly Pro Pro Gly Val Gly
 35 40 45

Lys Thr Ser Leu Ala Leu Ala Leu Ala Asn Thr Met Gly Trp Glu Ala
 50 55 60

Val Glu Leu Asn Ala Ser Asp Gln Arg Ser Trp Arg Val Ile Glu Arg
 65 70 75 80

Ile Val Gly Glu Gly Ala Phe Asn Glu Thr Ile Ser Asp Glu Gly Glu
 85 90 95

Phe Leu Ser Ser Arg Ile Gly Lys Leu Lys Leu Ile Ile Leu Asp Glu
 100 105 110

Val Asp Asn Ile His Lys Lys Glu Asp Val Gly Gly Glu Ala Ala Leu
 115 120 125

Ile Arg Leu Ile Lys Arg Lys Pro Ala Gln Pro Leu Ile Leu Ile Ala
 130 135 140

Asn Asp Pro Tyr Lys Leu Ser Pro Glu Leu Arg Asn Leu Cys Glu Met
 145 150 155 160

Ile Asn Phe Lys Arg Leu Thr Lys Gln Gln Val Ala Arg Val Leu Glu
165 170 175

Arg Ile Ala Leu Lys Glu Gly Ile Lys Val Asp Lys Ser Val Leu Leu
180 185 190

Lys Ile Ala Glu Asn Ala Gly Gly Asp Leu Arg Ala Ala Ile Asn Asp
195 200 205

Phe Gln Ala Leu Ala Glu Gly Lys Glu Glu Leu Lys Pro Glu Asp Val
210 215 220

Phe Leu Thr Lys Arg Thr Gln Glu Lys Asp Ile Phe Arg Val Met Gln
225 230 235 240

Met Ile Phe Lys Thr Lys Asn Pro Ala Val Tyr Asn Glu Ala Met Leu
245 250 255

Leu Asp Glu Ser Pro Glu Asp Val Ile His Trp Val Asp Glu Asn Leu
260 265 270

Pro Leu Glu Tyr Ser Gly Val Glu Leu Val Asn Ala Tyr Glu Ala Leu
275 280 285

Ser Arg Ala Asp Ile Phe Leu Gly Arg Val Arg Arg Arg Gln Phe Tyr
290 295 300

Arg Leu Trp Lys Tyr Ala Ser Tyr Leu Met Thr Val Gly Val Gln Gln
305 310 315 320

Met Lys Glu Glu Pro Lys Lys Gly Phe Thr Arg Tyr Arg Arg Pro Ala
325 330 335

Val Trp Gln Met Leu Phe Gln Leu Arg Gln Lys Arg Glu Met Thr Arg
340 345 350

Lys Ile Leu Glu Lys Ile Gly Lys Tyr Ser His Leu Ser Met Arg Lys
355 360 365

Ala Arg Thr Glu Met Phe Pro Val Ile Lys Leu Leu Leu Lys Glu Leu
370 375 380

Asp Val Asp Lys Ala Ala Thr Ile Ala Ala Phe Tyr Glu Phe Thr Lys
385 390 395 400

Glu Glu Leu Glu Phe Leu Val Gly Glu Lys Gly Asp Glu Ile Trp Lys
405 410 415

Tyr Val Glu Lys His Gly Met His Arg Ile Glu Asp Glu Thr Phe Leu

420 425 430
 Glu Ser Phe Val Lys Ala Glu Lys Glu Glu Lys Glu Glu Ser Val Glu
 435 440 445
 Glu Val Ala Glu Glu Lys Pro Glu Glu Glu Arg Glu Glu Pro Arg Ala
 450 455 460
 Arg Lys Lys Ala Gly Lys Asn Leu Thr Leu Asp Ser Phe Phe Ser
 465 470 475

(2) ANGABEN ZU SEQ ID NO: 8:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 516 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: Methanococcus jannaschii

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 8:

Met Leu Ser Trp Val Glu Lys Tyr Arg Pro Lys Ser Leu Lys Asp Val
 1 5 10 15
 Ala Gly His Glu Lys Val Lys Glu Lys Leu Lys Thr Trp Ile Glu Ser
 20 25 30
 Tyr Leu Lys Gly Glu Thr Pro Lys Pro Ile Leu Leu Val Gly Pro Pro
 35 40 45
 Gly Cys Gly Lys Thr Thr Leu Ala Tyr Ala Leu Ala Asn Asp Tyr Gly
 50 55 60
 Phe Glu Val Ile Glu Leu Asn Ala Ser Asp Lys Arg Asn Ser Ser Ala
 65 70 75 80
 Ile Lys Lys Val Val Gly His Ala Ala Thr Ser Ser Ser Ile Phe Gly
 85 90 95
 Lys Lys Phe Leu Ile Val Leu Asp Glu Val Asp Gly Ile Ser Gly Lys
 100 105 110

Glu Asp Ala Gly Gly Val Ser Glu Leu Ile Lys Val Ile Lys Lys Ala
 115 120 125

Lys Asn Pro Ile Ile Leu Thr Ala Asn Asp Ala Tyr Ala Pro Ser Ile
 130 135 140

Arg Ser Leu Leu Pro Tyr Val Glu Val Ile Gln Leu Asn Pro Val His
 145 150 155 160

Thr Asn Ser Val Tyr Lys Val Leu Lys Lys Ile Ala Glu Lys Glu Gly
 165 170 175

Leu Asp Val Asp Asp Lys Thr Leu Lys Met Ile Ala Gln His Ser Ala
 180 185 190

Gly Asp Leu Arg Ser Ala Ile Asn Asp Leu Glu Ala Leu Ala Leu Ser
 195 200 205

Gly Asp Leu Ser Tyr Glu Ala Ala Gln Lys Leu Pro Asp Arg Lys Arg
 210 215 220

Glu Ala Asn Ile Phe Asp Ala Leu Arg Val Ile Leu Lys Thr Thr His
 225 230 235 240

Tyr Gly Ile Ala Thr Thr Ala Leu Met Asn Val Asp Glu Thr Pro Asp
 245 250 255

Val Val Ile Glu Trp Ile Ala Glu Asn Val Pro Lys Glu Tyr Glu Lys
 260 265 270

Pro Glu Glu Val Ala Arg Ala Phe Glu Tyr Leu Ser Lys Ala Asp Arg
 275 280 285

Tyr Leu Gly Arg Val Met Arg Arg Gln Asn Tyr Ser Phe Trp Lys Tyr
 290 295 300

Ala Thr Thr Leu Met Thr Ala Gly Val Ala Leu Ser Lys Asp Glu Lys
 305 310 315 320

Tyr Arg Lys Trp Thr Pro Tyr Ser Tyr Pro Lys Ile Phe Arg Leu Leu
 325 330 335

Thr Lys Thr Lys Ala Glu Arg Glu Ile Leu Asn Lys Ile Leu Lys Lys
 340 345 350

Ile Gly Glu Lys Thr His Thr Ser Ser Lys Arg Ala Arg Phe Asp Leu
 355 360 365

Gln Met Leu Lys Leu Leu Ala Lys Glu Asn Pro Ser Val Ala Ala Asp

370 375 380
 Leu Val Asp Tyr Phe Glu Ile Lys Glu Asp Glu Leu Lys Val Leu Val
 385 390 395 400
 Gly Asp Lys Leu Ala Ser Glu Ile Leu Lys Ile Leu Lys Glu Lys Lys
 405 410 415
 Lys Leu Glu Arg Lys Lys Lys Lys Glu Lys Glu Lys Leu Glu Lys Glu
 420 425 430
 Lys Lys Lys Glu Glu Lys Ala Lys Glu Lys Gln Ser Asn Leu Ile Ile
 435 440 445
 Gln Pro Lys Glu Ile Lys Glu Glu Val Lys Ala Glu Val Glu Lys Lys
 450 455 460
 Glu Glu Val Lys Glu Lys Ile Val Glu Lys Pro Lys Ala Glu Glu Val
 465 470 475 480
 Lys Glu Lys Ser Lys Thr Glu Glu Lys Glu Thr Lys Lys Asp Lys Lys
 485 490 495
 Lys Gly Lys Lys Lys Lys Glu Asp Lys Gly Lys Gln Leu Thr Leu Asp
 500 505 510
 Ala Phe Phe Lys
 515

(2) ANGABEN ZU SEQ ID NO: 9:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 468 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: *Pyrococcus horikoshii*

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 9:

Met Pro Asp Val Pro Trp Ile Glu Lys Tyr Arg Pro Arg Lys Leu Ser
 1 5 10 15

Glu Ile Val Asn Gln Glu Gln Ala Leu Glu Lys Val Arg Ala Trp Ile
 20 25 30

Glu Ser Trp Leu His Gly Asn Pro Pro Lys Lys Lys Ala Leu Leu Leu
 35 40 45

Ala Gly Pro Pro Gly Ser Gly Lys Thr Thr Thr Val Tyr Ala Leu Ala
 50 55 60

His Glu Tyr Asn Phe Glu Val Ile Glu Leu Asn Ala Ser Asp Glu Arg
 65 70 75 80

Thr Tyr Asn Lys Ile Ala Arg Tyr Val Gln Ala Ala Tyr Thr Met Asp
 85 90 95

Ile Met Gly Lys Arg Arg Lys Ile Ile Phe Leu Asp Glu Ala Asp Asn
 100 105 110

Ile Glu Pro Ser Gly Ala Pro Glu Ile Ala Lys Leu Ile Asp Lys Ala
 115 120 125

Arg Asn Pro Ile Ile Met Ala Ala Asn His Tyr Trp Glu Val Pro Lys
 130 135 140

Glu Ile Arg Asp Arg Ala Glu Leu Val Glu Tyr Lys Arg Leu Asn Gln
 145 150 155 160

Arg Asp Val Ile Ser Ala Leu Val Arg Ile Leu Lys Arg Glu Gly Ile
 165 170 175

Thr Val Pro Lys Glu Ile Leu Thr Glu Ile Ala Lys Arg Ser Ser Gly
 180 185 190

Asp Leu Arg Ala Ala Ile Asn Asp Leu Gln Thr Ile Val Ala Gly Gly
 195 200 205

Tyr Glu Asp Ala Lys Tyr Val Leu Ala Tyr Arg Asp Val Glu Lys Thr
 210 215 220

Val Phe Gln Ser Leu Gly Met Val Phe Ser Ser Asp Asn Ala Lys Arg
 225 230 235 240

Ala Lys Leu Ala Leu Met Asn Leu Asp Met Ser Pro Asp Glu Phe Leu
 245 250 255

Leu Trp Val Asp Glu Asn Ile Pro His Met Tyr Leu Lys Pro Glu Glu
 260 265 270

Met Ala Arg Ala Tyr Glu Ala Ile Ser Arg Ala Asp Ile Tyr Leu Gly

275 280 285
 Arg Ala Gln Arg Thr Gly Asn Tyr Ser Leu Trp Lys Tyr Ala Ile Asp
 290 295 300
 Met Met Thr Ala Gly Val Ala Val Ala Gly Thr Lys Lys Lys Gly Phe
 305 310 315 320
 Ala Lys Phe Tyr Pro Pro Asn Thr Leu Lys Met Leu Ala Glu Ser Lys
 325 330 335
 Glu Glu Arg Ser Ile Arg Asp Ser Ile Ile Lys Lys Ile Met Lys Glu
 340 345 350
 Met His Met Ser Lys Leu Glu Ala Leu Glu Thr Met Lys Ile Leu Arg
 355 360 365
 Thr Ile Phe Glu Asn Asn Leu Asp Leu Ala Ala His Phe Thr Val Phe
 370 375 380
 Leu Glu Leu Thr Glu Lys Glu Val Glu Phe Leu Ala Gly Lys Glu Lys
 385 390 395 400
 Ala Gly Thr Ile Trp Gly Lys Thr Leu Ser Ile Arg Arg Arg Ile Lys
 405 410 415
 Glu Thr Glu Lys Ile Glu Glu Lys Ala Val Glu Glu Lys Val Glu Glu
 420 425 430
 Glu Glu Ala Glu Glu Glu Glu Glu Glu Arg Lys Glu Glu Glu Lys
 435 440 445
 Pro Lys Ala Glu Lys Lys Lys Gly Lys Gln Val Thr Leu Phe Asp Phe
 450 455 460
 Ile Lys Lys Asn
 465

(2) ANGABEN ZU SEQ ID NO: 10:

(i) SEQUENZKENNZEICHEN:

(A) LÄNGE: 479 Aminosäuren

(B) ART: Aminosäure

(C) STRANGFORM: Einzelstrang

(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

(A) ORGANISMUS: *Methanobacterium thermoautotrophicum*

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 10:

Met Ser Trp Thr Glu Lys Tyr Arg Pro Gly Ser Phe Asp Glu Val Val
1 5 10 15

Gly Asn Gln Lys Val Ile Ala Glu Ile Lys Glu Trp Ile Lys Ala Trp
20 25 30

Lys Ala Gly Lys Pro Gln Lys Pro Leu Leu Leu Val Gly Pro Pro Gly
35 40 45

Thr Gly Lys Thr Thr Leu Ala His Ile Ile Gly Lys Glu Phe Ser Asp
50 55 60

Thr Leu Glu Leu Asn Ala Ser Asp Arg Arg Ser Gln Asp Ala Leu Met
65 70 75 80

Arg Ser Ala Gly Glu Ala Ser Ala Thr Arg Ser Leu Phe Asn His Asp
85 90 95

Leu Lys Leu Ile Ile Leu Asp Glu Val Asp Gly Ile His Gly Asn Glu
100 105 110

Asp Arg Gly Gly Val Gln Ala Ile Asn Arg Ile Ile Lys Glu Ser Arg
115 120 125

His Pro Met Val Leu Thr Ala Asn Asp Pro Tyr Ser Lys Arg Leu Gln
130 135 140

Ser Ile Lys Pro Arg Cys Arg Val Leu Asn Leu Arg Lys Val His Thr
145 150 155 160

Ser Ser Ile Ala Ala Leu Arg Arg Ile Cys Arg Ala Glu Gly Ile
165 170 175

Glu Cys Pro Asp Asp Val Leu Arg Glu Leu Ala Lys Arg Ser Arg Gly
180 185 190

Asp Leu Arg Ser Ala Ile Asn Asp Leu Glu Ala Met Ala Glu Gly Glu
195 200 205

Glu Arg Ile Gly Glu Glu Leu Leu Lys Met Gly Glu Lys Asp Ala Thr
210 215 220

Ser Asn Leu Phe Asp Ala Val Arg Ala Val Leu Lys Ser Arg Asp Val

225 230 235 240
 Ser Lys Val Arg Glu Ala Met Arg Val Asp Asp Asp Pro Thr Leu Val
 245 250 255
 Leu Glu Phe Ile Ala Glu Asn Val Pro Arg Glu Tyr Glu Lys Pro Asn
 260 265 270
 Glu Ile Ser Arg Ala Tyr Asp Met Leu Ser Arg Ala Asp Ile Phe Phe
 275 280 285
 Gly Arg Ala Val Arg Thr Arg Asn Tyr Thr Tyr Trp Arg Tyr Ala Ser
 290 295 300
 Glu Leu Met Gly Pro Gly Val Ala Leu Ala Lys Asp Lys Thr Tyr Arg
 305 310 315 320
 Lys Phe Val Arg Tyr Thr Gly Ser Ser Ser Phe Arg Ile Leu Gly Lys
 325 330 335
 Thr Arg Lys Gln Arg Ser Leu Arg Asp Ser Val Ala Ala Lys Met Ala
 340 345 350
 Gly Lys Met His Ile Ser Pro Lys Val Ala Ile Ser Met Phe Pro Tyr
 355 360 365
 Met Glu Ile Leu Phe Glu Asn Asp Glu Met Ala Tyr Asp Ile Ser Glu
 370 375 380
 Phe Leu Glu Leu Arg Asp Glu Glu Ile Lys Leu Phe Arg Lys Arg Lys
 385 390 395 400
 Ile Lys Ala Pro Lys Arg Lys Lys Thr Pro Arg Lys Ala Glu Ile Lys
 405 410 415
 Val Gly Pro Leu Tyr Ser Gln Lys Lys Asp Lys Gly Ala Asp Lys Ser
 420 425 430
 Ile Asn Asp Lys Ala Thr Asp Lys Ser Ala Lys Thr Pro Ile Lys Ser
 435 440 445
 Ser Lys Lys Asp Asp Arg Pro Arg Asp Glu Ser Ser Ser Ser Asp
 450 455 460
 Asp Lys Lys Pro Lys Glu Lys Gln Thr Ser Leu Phe Gln Phe Ser
 465 470 475

(2) ANGABEN ZU SEQ ID NO: 11:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 261 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: Homo sapiens

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 11:

Met Phe Glu Ala Arg Leu Val Gln Gly Ser Ile Leu Lys Lys Val Leu
1 5 10 15

Glu Ala Leu Lys Asp Leu Ile Asn Glu Ala Cys Trp Asp Ile Ser Ser
20 25 30

Ser Gly Val Asn Leu Gln Ser Met Asp Ser Ser His Val Ser Leu Val
35 40 45

Gln Leu Thr Leu Arg Ser Glu Gly Phe Asp Thr Tyr Arg Cys Asp Arg
50 55 60

Asn Leu Ala Met Gly Val Asn Leu Thr Ser Met Ser Lys Ile Leu Lys
65 70 75 80

Cys Ala Gly Asn Glu Asp Ile Ile Thr Leu Arg Ala Glu Asp Asn Ala
85 90 95

Asp Thr Leu Ala Leu Val Phe Glu Ala Pro Asn Gln Glu Lys Val Ser
100 105 110

Asp Tyr Glu Met Lys Leu Met Asp Leu Asp Val Glu Gln Leu Gly Ile
115 120 125

Pro Glu Gln Glu Tyr Ser Cys Val Val Lys Met Pro Ser Gly Glu Phe
130 135 140

Ala Arg Ile Cys Arg Asp Leu Ser His Ile Gly Asp Ala Val Val Ile
145 150 155 160

Ser Cys Ala Lys Asp Gly Val Lys Phe Ser Ala Ser Gly Glu Leu Gly
165 170 175

Asn Gly Asn Ile Lys Leu Ser Gln Thr Ser Asn Val Asp Lys Glu Glu

180 185 190
 Glu Ala Val Thr Ile Glu Met Asn Glu Pro Val Gln Leu Thr Phe Ala
 195 200 205
 Leu Arg Tyr Leu Asn Phe Phe Thr Lys Ala Thr Pro Leu Ser Ser Thr
 210 215 220
 Val Thr Leu Ser Met Ser Ala Asp Val Pro Leu Val Val Glu Tyr Lys
 225 230 235 240
 Ile Ala Asp Met Gly His Leu Lys Tyr Tyr Leu Ala Pro Lys Ile Glu
 245 250 255
 Asp Glu Glu Gly Ser
 260

(2) ANGABEN ZU SEQ ID NO: 12:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 245 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: Archaeoglobus fulgidus

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 12:

Met Ile Asp Val Ile Met Thr Gly Glu Leu Leu Lys Thr Val Thr Arg
 1 5 10 15
 Ala Ile Val Ala Leu Val Ser Glu Ala Arg Ile His Phe Leu Glu Lys
 20 25 30
 Gly Leu His Ser Arg Ala Val Asp Pro Ala Asn Val Ala Met Val Ile
 35 40 45
 Val Asp Ile Pro Lys Asp Ser Phe Glu Val Tyr Asn Ile Asp Glu Glu
 50 55 60
 Lys Thr Ile Gly Val Asp Met Asp Arg Ile Phe Asp Ile Ser Lys Ser
 65 70 75 80

Ile Ser Thr Lys Asp Leu Val Glu Leu Ile Val Glu Asp Glu Ser Thr
 85 90 95
 Leu Lys Val Lys Phe Gly Ser Val Glu Tyr Lys Val Ala Leu Ile Asp
 100 105 110
 Pro Ser Ala Ile Arg Lys Glu Pro Arg Ile Pro Glu Leu Glu Leu Pro
 115 120 125
 Ala Lys Ile Val Met Asp Ala Gly Glu Phe Lys Lys Ala Ile Ala Ala
 130 135 140
 Ala Asp Lys Ile Ser Asp Gln Val Ile Phe Arg Ser Asp Lys Glu Gly
 145 150 155 160
 Phe Arg Ile Glu Ala Lys Gly Asp Val Asp Ser Ile Val Phe His Met
 165 170 175
 Thr Glu Thr Glu Leu Ile Glu Phe Asn Gly Gly Glu Ala Arg Ser Met
 180 185 190
 Phe Ser Val Asp Tyr Leu Lys Glu Phe Cys Lys Val Ala Gly Ser Gly
 195 200 205
 Asp Leu Leu Thr Ile His Leu Gly Thr Asn Tyr Pro Val Arg Leu Val
 210 215 220
 Phe Glu Leu Val Gly Gly Arg Ala Lys Val Glu Tyr Ile Leu Ala Pro
 225 230 235 240
 Arg Ile Glu Ser Glu
 245

(2) ANGABEN ZU SEQ ID NO: 13:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 247 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: *Methanococcus jannaschii*

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 13:

Met Phe Arg Gly Val Met Glu Ser Ala Lys Glu Phe Lys Lys Val Val
 1 5 10 15
 Asp Thr Ile Ser Thr Leu Leu Asp Glu Ile Cys Phe Glu Val Asp Glu
 20 25 30
 Glu Gly Ile Lys Ala Ser Ala Met Asp Pro Ser His Val Ala Leu Val
 35 40 45
 Ser Leu Glu Ile Pro Arg Leu Ala Phe Glu Glu Tyr Glu Ala Asp Ser
 50 55 60
 His Asp Ile Gly Ile Asp Leu Glu Ala Phe Lys Lys Val Met Asn Arg
 65 70 75 80
 Ala Lys Ala Lys Asp Arg Leu Ile Leu Glu Leu Asp Glu Glu Lys Asn
 85 90 95
 Lys Leu Asn Val Ile Phe Glu Asn Thr Gly Lys Arg Lys Phe Ser Leu
 100 105 110
 Ala Leu Leu Asp Ile Ser Ala Ser Ser Val Lys Val Pro Glu Ile Glu
 115 120 125
 Tyr Pro Asn Val Ile Met Ile Lys Gly Asp Ala Phe Lys Glu Ala Leu
 130 135 140
 Lys Asp Ala Asp Leu Phe Ser Asp Tyr Val Ile Leu Lys Ala Asp Glu
 145 150 155 160
 Asp Lys Phe Val Ile His Ala Lys Gly Asp Leu Asn Glu Asn Glu Ala
 165 170 175
 Ile Phe Glu Lys Asp Ser Ser Ala Ile Ile Ser Leu Glu Val Lys Glu
 180 185 190
 Glu Ala Lys Ser Ala Phe Asn Leu Asp Tyr Leu Met Asp Met Val Lys
 195 200 205
 Gly Val Ser Ser Gly Asp Ile Ile Lys Ile Tyr Leu Gly Asn Asp Met
 210 215 220
 Pro Leu Lys Leu Glu Tyr Ser Ile Ala Gly Val Asn Leu Thr Phe Leu
 225 230 235 240
 Leu Ala Pro Arg Ile Glu Gly
 245

(2) ANGABEN ZU SEQ ID NO: 14:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 249 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: Pyrococcus horikoshii

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 14:

Met Pro Phe Glu Ile Val Phe Glu Gly Ala Lys Glu Phe Ala Gln Leu
 1 5 10 15

Ile Glu Thr Ala Ser Arg Leu Ile Asp Glu Ala Ala Phe Lys Val Thr
 20 25 30

Glu Glu Gly Ile Ser Met Arg Ala Met Asp Pro Ser Arg Val Val Leu
 35 40 45

Ile Asp Leu Asn Leu Pro Ser Ser Ile Phe Ser Lys Tyr Glu Val Asp
 50 55 60

Gly Glu Glu Thr Ile Gly Val Asn Met Asp His Leu Lys Lys Val Leu
 65 70 75 80

Lys Arg Gly Lys Ala Lys Asp Thr Leu Ile Leu Arg Lys Gly Glu Glu
 85 90 95

Asn Phe Leu Glu Ile Ser Leu Gln Gly Thr Ala Thr Arg Thr Phe Arg
 100 105 110

Leu Pro Leu Ile Asp Val Glu Glu Ile Glu Val Glu Leu Pro Asp Leu
 115 120 125

Pro Tyr Thr Ala Lys Val Val Val Leu Gly Glu Val Leu Lys Glu Ala
 130 135 140

Val Lys Asp Ala Ser Leu Val Ser Asp Ser Ile Lys Phe Met Ala Lys
 145 150 155 160

Glu Asn Glu Phe Ile Met Arg Ala Glu Gly Glu Thr Gln Glu Val Glu
 165 170 175

Val Lys Leu Thr Leu Glu Asp Glu Gly Leu Leu Asp Ile Glu Val Gln

180 185 190
 Glu Glu Thr Lys Ser Ala Tyr Gly Val Ser Tyr Leu Ala Asp Met Val
 195 200 205
 Lys Gly Ile Gly Lys Ala Asp Glu Val Thr Met Arg Phe Gly Asn Glu
 210 215 220
 Met Pro Met Gln Met Glu Tyr Tyr Ile Arg Asp Glu Gly Arg Leu Thr
 225 230 235 240
 Phe Leu Leu Ala Pro Arg Val Glu Glu
 245

(2) ANGABEN ZU SEQ ID NO: 15:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 244 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: Methanobacterium thermoautotrophicum

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 15:

Met Phe Lys Ala Glu Leu Asn Asp Pro Asn Ile Leu Arg Thr Ser Phe
 1 5 10 15
 Asp Ala Ile Ser Ser Ile Val Asp Glu Val Gln Ile Gln Leu Ser Ala
 20 25 30
 Glu Gly Leu Arg Leu Asp Ala Leu Asp Arg Ser His Ile Thr Tyr Val
 35 40 45
 His Leu Glu Leu Lys Ala Glu Leu Phe Asp Glu Tyr Val Cys Asp Glu
 50 55 60
 Pro Glu Arg Ile Asn Val Asp Thr Glu Glu Leu Met Lys Val Leu Lys
 65 70 75 80
 Arg Ala Lys Ala Asn Asp Arg Val Ile Leu Ser Thr Asp Glu Gly Asn
 85 90 95

Leu Ile Ile Gln Phe Glu Gly Glu Ala Val Arg Thr Phe Lys Ile Arg
 100 105 110

Leu Ile Asp Ile Glu Tyr Glu Thr Pro Ser Pro Pro Glu Ile Glu Tyr
 115 120 125

Glu Asn Glu Phe Glu Val Pro Phe Gln Leu Leu Lys Asp Ser Ile Ala
 130 135 140

Asp Ile Asp Ile Phe Ser Asp Lys Ile Thr Phe Arg Val Asp Glu Asp
 145 150 155 160

Arg Phe Ile Ala Ser Ala Glu Gly Glu Phe Gly Asp Ala Gln Ile Glu
 165 170 175

Tyr Leu His Gly Glu Arg Ile Asp Lys Pro Ala Arg Ser Ile Tyr Ser
 180 185 190

Leu Asp Lys Ile Lys Glu Met Leu Lys Ala Asp Lys Phe Ser Glu Thr
 195 200 205

Ala Ile Ile Asn Leu Gly Asp Asp Met Pro Leu Lys Leu Thr Leu Lys
 210 215 220

Met Ala Ser Lys Glu Gly Glu Leu Ser Phe Leu Leu Ala Pro Arg Ile
 225 230 235 240

Glu Ala Glu Glu

(2) ANGABEN ZU SEQ ID NO: 16:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 469 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: Homo sapiens

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 16:

Met Phe Ser Glu Gln Ala Ala Gln Arg Ala His Thr Leu Leu Ser Pro
 1 5 10 15

Pro Ser Ala Asn Asn Ala Thr Phe Ala Arg Val Pro Val Ala Thr Tyr
 20 25 30
 Thr Asn Ser Ser Gln Pro Phe Arg Leu Gly Glu Arg Ser Phe Ser Arg
 35 40 45
 Gln Tyr Ala His Ile Tyr Ala Thr Arg Leu Ile Gln Met Arg Pro Phe
 50 55 60
 Leu Glu Asn Arg Ala Gln Gln His Trp Gly Ser Gly Val Gly Val Lys
 65 70 75 80
 Lys Leu Cys Glu Leu Gln Pro Glu Glu Lys Cys Cys Val Val Gly Thr
 85 90 95
 Leu Phe Lys Ala Met Pro Leu Gln Pro Ser Ile Leu Arg Glu Val Ser
 100 105 110
 Glu Glu His Asn Leu Leu Pro Gln Pro Pro Arg Ser Lys Tyr Ile His
 115 120 125
 Pro Asp Asp Glu Leu Val Leu Glu Asp Glu Leu Gln Arg Ile Lys Leu
 130 135 140
 Lys Gly Thr Ile Asp Val Ser Lys Leu Val Thr Gly Thr Val Leu Ala
 145 150 155 160
 Val Phe Gly Ser Val Arg Asp Asp Gly Lys Phe Leu Val Glu Asp Tyr
 165 170 175
 Cys Phe Ala Asp Leu Ala Pro Gln Lys Pro Ala Pro Pro Leu Asp Thr
 180 185 190
 Asp Arg Phe Val Leu Leu Val Ser Gly Leu Gly Leu Gly Gly Gly Gly
 195 200 205
 Gly Glu Ser Leu Leu Gly Thr Gln Leu Leu Val Asp Val Val Thr Gly
 210 215 220
 Gln Leu Gly Asp Glu Gly Glu Gln Cys Ser Ala Ala His Val Ser Arg
 225 230 235 240
 Val Ile Leu Ala Gly Asn Leu Leu Ser His Ser Thr Gln Ser Arg Asp
 245 250 255
 Ser Ile Asn Lys Ala Lys Tyr Leu Thr Lys Lys Thr Gln Ala Ala Ser
 260 265 270
 Val Glu Ala Val Lys Met Leu Asp Glu Ile Leu Leu Gln Leu Ser Ala

275 280 285
 Ser Val Pro Val Asp Val Met Pro Gly Glu Phe Asp Pro Thr Asn Tyr
 290 295 300
 Thr Leu Pro Gln Gln Pro Leu His Pro Cys Met Phe Pro Leu Ala Thr
 305 310 315 320
 Ala Tyr Ser Thr Leu Gln Leu Val Thr Asn Pro Tyr Gln Ala Thr Ile
 325 330 335
 Asp Gly Val Arg Phe Leu Gly Thr Ser Gly Gln Asn Val Ser Asp Ile
 340 345 350
 Phe Arg Tyr Ser Ser Met Glu Asp His Leu Glu Ile Leu Glu Trp Thr
 355 360 365
 Leu Arg Val Arg His Ile Ser Pro Thr Ala Pro Asp Thr Leu Gly Cys
 370 375 380
 Tyr Pro Phe Tyr Lys Thr Asp Pro Phe Ile Phe Pro Glu Cys Pro His
 385 390 395 400
 Val Tyr Phe Cys Gly Asn Thr Pro Ser Phe Gly Ser Lys Ile Ile Arg
 405 410 415
 Gly Pro Glu Asp Gln Thr Val Leu Leu Val Thr Val Pro Asp Phe Ser
 420 425 430
 Ala Thr Gln Thr Ala Cys Leu Val Asn Leu Arg Ser Leu Ala Cys Gln
 435 440 445
 Pro Ile Ser Phe Ser Gly Phe Gly Ala Glu Asp Asp Asp Leu Gly Gly
 450 455 460
 Leu Gly Leu Gly Pro
 465

(2) ANGABEN ZU SEQ ID NO: 17:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 488 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

(A) ORGANISMUS: Archaeoglobus fulgidus

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 17:

Met Val Ile Lys Asn Ile Asp Ala Ala Thr Val Ala Lys Lys Phe Leu
1 5 10 15

Val Arg Gly Tyr Asn Ile Asp Pro Lys Ala Ala Glu Leu Ile Cys Lys
20 25 30

Ser Gly Leu Phe Ser Asp Glu Leu Val Asp Lys Ile Cys Arg Ile Ala
35 40 45

Asn Gly Gly Phe Ile Ile Glu Lys Ser Val Val Glu Glu Phe Leu Arg
50 55 60

Asn Leu Ser Asn Leu Lys Pro Ala Thr Leu Thr Pro Arg Pro Glu Glu
65 70 75 80

Arg Lys Val Glu Glu Val Lys Ala Ser Cys Ile Ala Leu Lys Val Ile
85 90 95

Lys Asp Ile Thr Gly Lys Ser Ser Cys Gln Gly Asn Val Glu Asp Phe
100 105 110

Leu Met Tyr Phe Asn Ser Arg Leu Glu Lys Leu Ser Arg Ile Ile Arg
115 120 125

Ser Arg Val Asn Thr Thr Pro Ile Ala His Ala Gly Lys Val Arg Gly
130 135 140

Asn Val Ser Val Val Gly Met Val Asn Glu Val Tyr Glu Arg Gly Asp
145 150 155 160

Lys Cys Tyr Ile Arg Leu Glu Asp Thr Thr Gly Thr Ile Thr Cys Val
165 170 175

Ala Thr Gly Lys Asn Ala Glu Val Ala Arg Glu Leu Leu Gly Asp Glu
180 185 190

Val Ile Gly Val Thr Gly Leu Leu Lys Gly Ser Ser Leu Tyr Ala Asn
195 200 205

Arg Ile Val Phe Pro Asp Val Pro Ile Asn Gly Asn Gly Glu Lys Lys
210 215 220

Arg Asp Phe Tyr Ile Val Phe Leu Ser Asp Thr His Phe Gly Ser Lys

225 230 235 240
 Glu Phe Leu Glu Lys Glu Trp Glu Met Phe Val Arg Trp Leu Lys Gly
 245 250 255
 Glu Val Gly Gly Lys Lys Ser Gln Asn Leu Ala Glu Lys Val Lys Tyr
 260 265 270
 Ile Val Ile Ala Gly Asp Ile Val Asp Gly Ile Gly Val Tyr Pro Gly
 275 280 285
 Gln Glu Asp Asp Leu Ala Ile Ser Asp Ile Tyr Gly Gln Tyr Glu Phe
 290 295 300
 Ala Ala Ser His Leu Asp Glu Ile Pro Lys Glu Ile Lys Ile Ile Val
 305 310 315 320
 Ser Pro Gly Asn His Asp Ala Val Arg Gln Ala Glu Pro Gln Pro Ala
 325 330 335
 Phe Glu Gly Glu Ile Arg Ser Leu Phe Pro Lys Asn Val Glu His Val
 340 345 350
 Gly Asn Pro Ala Tyr Val Asp Ile Glu Gly Val Lys Val Leu Ile Tyr
 355 360 365
 His Gly Arg Ser Ile Asp Asp Ile Ile Ser Lys Ile Pro Arg Leu Ser
 370 375 380
 Tyr Asp Glu Pro Gln Lys Val Met Glu Glu Leu Leu Lys Arg Arg His
 385 390 395 400
 Leu Ser Pro Ile Tyr Gly Gly Arg Thr Pro Leu Ala Pro Glu Arg Glu
 405 410 415
 Asp Tyr Leu Val Ile Glu Asp Val Pro Asp Ile Leu His Cys Gly His
 420 425 430
 Ile His Thr Tyr Gly Thr Gly Phe Tyr Arg Gly Val Phe Met Val Asn
 435 440 445
 Ser Ser Thr Trp Gln Ala Gln Thr Glu Phe Gln Lys Lys Val Asn Leu
 450 455 460
 Asn Pro Met Pro Gly Asn Val Ala Val Tyr Arg Pro Gly Gly Glu Val
 465 470 475 480
 Ile Arg Leu Arg Phe Tyr Gly Glu
 485

(2) ANGABEN ZU SEQ ID NO: 18:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 594 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: *Methanococcus jannaschii*

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 18:

Met Glu Ile Ile Asn Lys Phe Leu Asp Leu Glu Ala Leu Leu Ser Pro
1 5 10 15

Thr Val Tyr Glu Lys Leu Lys Asn Phe Asp Glu Glu Lys Leu Lys Arg
 20 25 30

Leu Ile Gln Lys Ile Arg Glu Phe Lys Lys Tyr Asn Asn Ala Phe Ile
 35 40 45

Leu Leu Asp Glu Lys Phe Leu Asp Ile Phe Leu Gln Lys Asp Leu Asp
 50 55 60

Glu Ile Ile Asn Glu Tyr Lys Asp Phe Asp Phe Ile Phe Tyr Tyr Thr
65 70 75 80

Gly Glu Glu Glu Lys Glu Lys Pro Lys Glu Val Lys Lys Glu Ile Lys
 85 90 95

Lys Glu Thr Glu Glu Lys Ile Glu Lys Glu Lys Ile Glu Phe Val Lys
 100 105 110

Lys Glu Glu Lys Glu Gln Phe Ile Lys Lys Ser Asp Glu Asp Val Glu
 115 120 125

Glu Lys Leu Lys Gln Leu Ile Ser Lys Glu Glu Lys Lys Glu Asp Phe
 130 135 140

Asp Ala Glu Arg Ala Lys Arg Tyr Glu His Ile Thr Lys Ile Lys Glu
145 150 155 160

Ser Val Asn Ser Arg Ile Lys Trp Ile Ala Lys Asp Ile Asp Ala Val
 165 170 175

Ile Glu Ile Tyr Glu Asp Ser Asp Val Ser Gly Lys Ser Thr Cys Thr
 180 185 190

Gly Thr Ile Glu Asp Phe Val Lys Tyr Phe Arg Asp Arg Phe Glu Arg
 195 200 205

Leu Lys Val Phe Ile Glu Arg Lys Ala Gln Arg Lys Gly Tyr Pro Leu
 210 215 220

Lys Asp Ile Lys Lys Met Lys Gly Gln Lys Asp Ile Phe Val Val Gly
 225 230 235 240

Ile Val Ser Asp Val Asp Ser Thr Arg Asn Gly Asn Leu Ile Val Arg
 245 250 255

Ile Glu Asp Thr Glu Asp Glu Ala Thr Leu Ile Leu Pro Lys Glu Lys
 260 265 270

Ile Glu Ala Gly Lys Ile Pro Asp Asp Ile Leu Leu Asp Glu Val Ile
 275 280 285

Gly Ala Ile Gly Thr Val Ser Lys Ser Gly Ser Ser Ile Tyr Val Asp
 290 295 300

Glu Ile Ile Arg Pro Ala Leu Pro Pro Lys Glu Pro Lys Arg Ile Asp
 305 310 315 320

Glu Glu Ile Tyr Met Ala Phe Leu Ser Asp Ile His Val Gly Ser Lys
 325 330 335

Glu Phe Leu His Lys Glu Phe Glu Lys Phe Ile Arg Phe Leu Asn Gly
 340 345 350

Asp Val Asp Asn Glu Leu Glu Glu Lys Val Val Ser Arg Leu Lys Tyr
 355 360 365

Ile Cys Ile Ala Gly Asp Leu Val Asp Gly Val Gly Val Tyr Pro Gly
 370 375 380

Gln Glu Glu Asp Leu Tyr Glu Val Asp Ile Ile Glu Gln Tyr Arg Glu
 385 390 395 400

Ile Ala Met Tyr Leu Asp Gln Ile Pro Glu His Ile Ser Ile Ile Ile
 405 410 415

Ser Pro Gly Asn His Asp Ala Val Arg Pro Ala Glu Pro Gln Pro Lys
 420 425 430

Leu Pro Glu Lys Ile Thr Lys Leu Phe Asn Arg Asp Asn Ile Tyr Phe

435 440 445
 Val Gly Asn Pro Cys Thr Leu Asn Ile His Gly Phe Asp Thr Leu Leu
 450 455 460
 Tyr His Gly Arg Ser Phe Asp Asp Leu Val Gly Gln Ile Arg Ala Ala
 465 470 475 480
 Ser Tyr Glu Asn Pro Val Thr Ile Met Lys Glu Leu Ile Lys Arg Arg
 485 490 495
 Leu Leu Cys Pro Thr Tyr Gly Gly Arg Cys Pro Ile Ala Pro Glu His
 500 505 510
 Lys Asp Tyr Leu Val Ile Asp Arg Asp Ile Asp Ile Leu His Thr Gly
 515 520 525
 His Ile His Ile Asn Gly Tyr Gly Ile Tyr Arg Gly Val Val Met Val
 530 535 540
 Asn Ser Gly Thr Phe Gln Glu Gln Thr Asp Phe Gln Lys Arg Met Gly
 545 550 555 560
 Ile Ser Pro Thr Pro Ala Ile Val Pro Ile Ile Asn Met Ala Lys Val
 565 570 575
 Gly Glu Lys Gly His Tyr Leu Glu Trp Asp Arg Gly Val Leu Glu Val
 580 585 590
 Arg Tyr

(2) ANGABEN ZU SEQ ID NO: 19:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 622 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: *Pyrococcus horikoshii*

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 19:

Met Asp Glu Phe Val Lys Gly Leu Met Lys Asn Gly Tyr Leu Ile Thr
 1 5 10 15
 Pro Ser Ala Tyr Tyr Leu Leu Val Gly His Phe Asn Glu Gly Lys Phe
 20 25 30
 Ser Leu Ile Glu Leu Ile Lys Phe Ala Lys Ser Arg Glu Thr Phe Ile
 35 40 45
 Ile Asp Asp Glu Ile Ala Asn Glu Phe Leu Lys Ser Ile Gly Ala Glu
 50 55 60
 Val Glu Leu Pro Gln Glu Ile Lys Glu Gly Tyr Ile Ser Thr Gly Glu
 65 70 75 80
 Gly Ser Gln Lys Val Pro Asp His Glu Glu Leu Glu Lys Ile Thr Asn
 85 90 95
 Glu Ser Ser Val Glu Ser Ser Ile Ser Thr Gly Glu Thr Pro Lys Thr
 100 105 110
 Glu Glu Leu Gln Pro Thr Leu Asp Ile Leu Glu Glu Glu Ile Gly Asp
 115 120 125
 Ile Glu Gly Gly Glu Ser Ser Ile Ser Thr Gly Asp Glu Val Pro Glu
 130 135 140
 Val Glu Asn Asn Asn Gly Gly Thr Val Val Val Phe Asp Lys Tyr Gly
 145 150 155 160
 Tyr Pro Phe Thr Tyr Val Pro Glu Glu Ile Glu Glu Glu Leu Glu Glu
 165 170 175
 Tyr Pro Lys Tyr Glu Asp Val Thr Ile Glu Ile Asn Pro Asn Leu Glu
 180 185 190
 Val Val Pro Ile Glu Lys Asp Tyr Glu Ile Lys Phe Asp Val Arg Arg
 195 200 205
 Val Lys Leu Lys Pro Pro Lys Val Lys Ser Gly Ser Gly Lys Glu Gly
 210 215 220
 Glu Ile Ile Val Glu Ala Tyr Ala Ser Leu Phe Arg Ser Arg Leu Arg
 225 230 235 240
 Lys Leu Arg Arg Ile Leu Arg Glu Asn Pro Glu Val Ser Asn Val Ile
 245 250 255
 Asp Ile Lys Lys Leu Lys Tyr Val Lys Gly Asp Glu Glu Val Thr Ile

260	265	270
Ile Gly Leu Val Asn Ser Lys Lys Glu Thr Ser Lys Gly Leu Ile Phe		
275	280	285
Glu Val Glu Asp Gln Thr Asp Arg Val Lys Val Phe Leu Pro Lys Asp		
290	295	300
Ser Glu Asp Tyr Arg Glu Ala Leu Lys Val Leu Pro Asp Ala Val Val		
305	310	315
		320
Ala Phe Lys Gly Val Tyr Ser Lys Arg Gly Ile Phe Phe Ala Asn Arg		
325	330	335
Phe Tyr Leu Pro Asp Val Pro Leu Tyr Arg Lys Gln Lys Pro Pro Leu		
340	345	350
Glu Glu Lys Val Tyr Ala Val Leu Thr Ser Asp Ile His Val Gly Ser		
355	360	365
Lys Glu Phe Cys Glu Lys Ala Phe Ile Lys Phe Leu Glu Trp Leu Asn		
370	375	380
Gly Tyr Val Glu Ser Lys Glu Glu Glu Glu Ile Val Ser Arg Ile Arg		
385	390	395
		400
Tyr Leu Ile Ile Ala Gly Asp Val Val Asp Gly Ile Gly Ile Tyr Pro		
405	410	415
Gly Gln Tyr Ser Asp Leu Ile Ile Pro Asp Ile Phe Asp Gln Tyr Glu		
420	425	430
Ala Leu Ala Asn Leu Leu Ser Asn Val Pro Lys His Ile Thr Ile Phe		
435	440	445
Ile Gly Pro Gly Asn His Asp Ala Ala Arg Pro Ala Ile Pro Gln Pro		
450	455	460
Glu Phe Tyr Glu Glu Tyr Ala Lys Pro Leu Tyr Lys Leu Lys Asn Thr		
465	470	475
		480
Val Ile Ile Ser Asn Pro Ala Val Ile Arg Leu His Gly Arg Asp Phe		
485	490	495
Leu Ile Ala His Gly Arg Gly Ile Glu Asp Val Val Ser Phe Val Pro		
500	505	510
Gly Leu Thr His His Lys Pro Gly Leu Pro Met Val Glu Leu Leu Lys		
515	520	525

Met Arg His Leu Ala Pro Thr Phe Gly Gly Lys Val Pro Ile Ala Pro
 530 535 540

Asp Pro Glu Asp Leu Leu Val Ile Glu Glu Val Pro Asp Leu Val Gln
 545 550 555 560

Met Gly His Val His Val Tyr Asp Thr Ala Val Tyr Arg Gly Val Gln
 565 570 575

Leu Val Asn Ser Ala Thr Trp Gln Ala Gln Thr Glu Phe Gln Lys Met
 580 585 590

Val Asn Ile Val Pro Thr Pro Gly Leu Val Pro Ile Val Asp Val Glu
 595 600 605

Ser Ala Arg Val Ile Lys Val Leu Asp Phe Ser Arg Trp Cys
 610 615 620

(2) ANGABEN ZU SEQ ID NO: 20:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 482 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: Methanobacterium thermoautotrophicum

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 20:

Met Asn Glu Ile Ile Gly Lys Phe Ala Arg Glu Gly Ile Leu Ile Glu
 1 5 10 15

Asp Asn Ala Tyr Phe Arg Leu Arg Glu Met Asp Asp Pro Ala Ser Val
 20 25 30

Ser Ser Glu Leu Ile Val Lys Ile Lys Ser Asn Gly Gly Lys Phe Thr
 35 40 45

Val Leu Thr Ser Glu Met Leu Asp Glu Phe Phe Glu Ile Asp Asn Pro
 50 55 60

Ala Glu Ile Lys Ala Arg Gly Pro Leu Met Val Pro Ala Glu Arg Asp
 65 70 75 80

Phe Asp Phe Glu Val Ile Ser Asp Thr Ser Asn Arg Ser Tyr Thr Ser
 85 90 95

Gly Glu Ile Gly Asp Met Ile Ala Tyr Phe Asn Ser Arg Tyr Ser Ser
 100 105 110

Leu Lys Asn Leu Leu Ser Lys Arg Pro Glu Leu Lys Gly His Ile Pro
 115 120 125

Ile Ala Asp Leu Arg Gly Gly Glu Asp Val Val Ser Ile Ile Gly Met
 130 135 140

Val Asn Asp Val Arg Asn Thr Lys Asn Asn His Arg Ile Ile Glu Leu
 145 150 155 160

Glu Asp Asp Thr Gly Glu Ile Ser Val Val Val His Asn Glu Asn His
 165 170 175

Lys Leu Phe Glu Lys Ser Glu Lys Ile Val Arg Asp Glu Val Val Gly
 180 185 190

Val His Gly Thr Lys Lys Gly Arg Phe Val Val Ala Ser Glu Ile Phe
 195 200 205

His Pro Gly Val Pro Arg Ile Gln Glu Lys Glu Met Asp Phe Ser Val
 210 215 220

Ala Phe Ile Ser Asp Val His Ile Gly Ser Gln Thr Phe Leu Glu Asp
 225 230 235 240

Ala Phe Met Lys Phe Val Lys Trp Ile Asn Gly Asp Phe Gly Ser Glu
 245 250 255

Glu Gln Arg Ser Leu Ala Ala Asp Val Lys Tyr Leu Val Val Ala Gly
 260 265 270

Asp Ile Val Asp Gly Ile Gly Ile Tyr Pro Gly Gln Glu Lys Glu Leu
 275 280 285

Leu Ile Arg Asp Ile His Glu Gln Tyr Glu Glu Ala Ala Arg Leu Phe
 290 295 300

Gly Asp Ile Arg Ser Asp Ile Lys Ile Val Met Ile Pro Gly Asn His
 305 310 315 320

Asp Ser Ser Arg Ile Ala Glu Pro Gln Pro Ala Ile Pro Glu Glu Tyr
 325 330 335

Ala Lys Ser Leu Tyr Ser Ile Arg Asn Ile Glu Phe Leu Ser Asn Pro

340 345 350
 Ser Leu Val Ser Leu Asp Gly Val Arg Thr Leu Ile Tyr His Gly Arg
 355 360 365
 Ser Phe Asp Asp Met Ala Met Ser Val Asn Gly Leu Ser His Glu Arg
 370 375 380
 Ser Asp Leu Ile Met Glu Glu Leu Leu Glu Lys Arg His Leu Ala Pro
 385 390 395 400
 Ile Tyr Gly Glu Arg Thr Pro Leu Ala Ser Glu Ile Glu Asp His Leu
 405 410 415
 Val Ile Asp Glu Val Pro His Val Leu His Thr Gly His Val His Ile
 420 425 430
 Asn Ala Tyr Lys Lys Tyr Lys Gly Val His Leu Ile Asn Ser Gly Thr
 435 440 445
 Phe Gln Ser Gln Thr Glu Phe Gln Lys Ile Tyr Asn Ile Val Pro Thr
 450 455 460
 Cys Gly Gln Val Pro Val Leu Asn Arg Gly Val Met Lys Leu Leu Glu
 465 470 475 480
 Phe Ser

(2) ANGABEN ZU SEQ ID NO: 21:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 613 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: *Pyrococcus furiosus*

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 21:

Met Asp Glu Phe Val Lys Ser Leu Leu Lys Ala Asn Tyr Leu Ile Thr
 1 5 10 15

Pro Ser Ala Tyr Tyr Leu Leu Arg Glu Tyr Tyr Glu Lys Gly Glu Phe
 20 25 30

Ser Ile Val Glu Leu Val Lys Phe Ala Arg Ser Arg Glu Ser Tyr Ile
 35 40 45

Ile Thr Asp Ala Leu Ala Thr Glu Phe Leu Lys Val Lys Gly Leu Glu
 50 55 60

Pro Ile Leu Pro Val Glu Thr Lys Gly Gly Phe Val Ser Thr Gly Glu
 65 70 75 80

Ser Gln Lys Glu Gln Ser Tyr Glu Glu Ser Phe Gly Thr Lys Glu Glu
 85 90 95

Ile Ser Gln Glu Ile Lys Glu Gly Glu Ser Phe Ile Ser Thr Gly Ser
 100 105 110

Glu Pro Leu Glu Glu Glu Leu Asn Ser Ile Gly Ile Glu Glu Ile Gly
 115 120 125

Ala Asn Glu Glu Leu Val Ser Asn Gly Asn Asp Asn Gly Gly Glu Ala
 130 135 140

Ile Val Phe Asp Lys Tyr Gly Tyr Pro Met Val Tyr Ala Pro Glu Glu
 145 150 155 160

Ile Glu Val Glu Glu Lys Glu Tyr Ser Lys Tyr Glu Asp Leu Thr Ile
 165 170 175

Pro Met Asn Pro Asp Phe Asn Tyr Val Glu Ile Lys Glu Asp Tyr Asp
 180 185 190

Val Val Phe Asp Val Arg Asn Val Lys Leu Lys Pro Pro Lys Val Lys
 195 200 205

Asn Gly Asn Gly Lys Glu Gly Glu Ile Ile Val Glu Ala Tyr Ala Ser
 210 215 220

Leu Phe Arg Ser Arg Leu Lys Lys Leu Arg Lys Ile Leu Arg Glu Asn
 225 230 235 240

Pro Glu Leu Asp Asn Val Val Asp Ile Gly Lys Leu Lys Tyr Val Lys
 245 250 255

Glu Asp Glu Thr Val Thr Ile Ile Gly Leu Val Asn Ser Lys Arg Glu
 260 265 270

Val Asn Lys Gly Leu Ile Phe Glu Ile Glu Asp Leu Thr Gly Lys Val

275 280 285
 Lys Val Phe Leu Pro Lys Asp Ser Glu Asp Tyr Arg Glu Ala Phe Lys
 290 295 300
 Val Leu Pro Asp Ala Val Val Ala Phe Lys Gly Val Tyr Ser Lys Arg
 305 310 315 320
 Gly Ile Leu Tyr Ala Asn Lys Phe Tyr Leu Pro Asp Val Pro Leu Tyr
 325 330 335
 Arg Arg Gln Lys Pro Pro Leu Glu Glu Lys Val Tyr Ala Ile Leu Ile
 340 345 350
 Ser Asp Ile His Val Gly Ser Lys Glu Phe Cys Glu Asn Ala Phe Ile
 355 360 365
 Lys Phe Leu Glu Trp Leu Asn Gly Asn Val Glu Thr Lys Glu Glu Glu
 370 375 380
 Glu Ile Val Ser Arg Val Lys Tyr Leu Ile Ile Ala Gly Asp Val Val
 385 390 395 400
 Asp Gly Val Gly Val Tyr Pro Gly Gln Tyr Ala Asp Leu Thr Ile Pro
 405 410 415
 Asp Ile Phe Asp Gln Tyr Glu Ala Leu Ala Asn Leu Leu Ser His Val
 420 425 430
 Pro Lys His Ile Thr Met Phe Ile Ala Pro Gly Asn His Asp Ala Ala
 435 440 445
 Arg Gln Ala Ile Pro Gln Pro Glu Phe Tyr Lys Glu Tyr Ala Lys Pro
 450 455 460
 Ile Tyr Lys Leu Lys Asn Ala Val Ile Ile Ser Asn Pro Ala Val Ile
 465 470 475 480
 Arg Leu His Gly Arg Asp Phe Leu Ile Ala His Gly Arg Gly Ile Glu
 485 490 495
 Asp Val Val Gly Ser Val Pro Gly Leu Thr His His Lys Pro Gly Leu
 500 505 510
 Pro Met Val Glu Leu Leu Lys Met Arg His Val Ala Pro Met Phe Gly
 515 520 525
 Gly Lys Val Pro Ile Ala Pro Asp Pro Glu Asp Leu Leu Val Ile Glu
 530 535 540

Glu Val Pro Asp Val Val His Met Gly His Val His Val Tyr Asp Ala
545 550 555 560

Val Val Tyr Arg Gly Val Gln Leu Val Asn Ser Ala Thr Trp Gln Ala
 565 570 575

Gln Thr Glu Phe Gln Lys Met Val Asn Ile Val Pro Thr Pro Ala Lys
 580 585 590

Val Pro Val Val Asp Ile Asp Thr Ala Lys Val Val Lys Val Leu Asp
 595 600 605

Phe Ser Gly Trp Cys
610

(2) ANGABEN ZU SEQ ID NO: 22:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 1107 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: Homo sapiens

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 22:

Met Asp Gly Lys Arg Arg Pro Gly Pro Gly Pro Gly Val Pro Pro Lys
1 5 10 15

Arg Ala Arg Gly Gly Leu Trp Asp Asp Asp Ala Pro Trp Pro Ser
 20 25 30

Gln Phe Glu Glu Asp Leu Ala Leu Met Glu Glu Met Glu Ala Glu His
 35 40 45

Arg Leu Gln Glu Gln Glu Glu Glu Leu Gln Ser Val Leu Glu Gly
 50 55 60

Val Ala Asp Gly Gln Val Pro Pro Ser Ala Ile Asp Pro Arg Trp Leu
65 70 75 80

Arg Pro Thr Pro Pro Ala Leu Asp Pro Gln Thr Glu Pro Leu Ile Phe
 85 90 95

Gln Gln Leu Glu Ile Asp His Tyr Val Gly Pro Ala Gln Pro Val Pro
100 105 110

Gly Gly Pro Pro Pro Ser Arg Gly Ser Val Pro Val Leu Arg Ala Phe
115 120 125

Gly Val Thr Asp Glu Gly Phe Ser Val Cys Cys His Ile His Gly Phe
130 135 140

Ala Pro Tyr Phe Tyr Thr Pro Ala Pro Pro Gly Phe Gly Pro Glu His
145 150 155 160

Met Gly Asp Leu Gln Arg Glu Leu Asn Leu Ala Ile Ser Arg Asp Ser
165 170 175

Arg Gly Gly Arg Glu Leu Thr Gly Pro Ala Val Leu Ala Val Glu Leu
180 185 190

Cys Ser Arg Glu Ser Met Phe Gly Tyr His Gly His Gly Pro Ser Pro
195 200 205

Phe Leu Arg Ile Thr Val Ala Leu Pro Arg Leu Val Ala Pro Ala Arg
210 215 220

Arg Leu Leu Glu Gln Gly Ile Arg Val Ala Gly Leu Gly Thr Pro Ser
225 230 235 240

Phe Ala Pro Tyr Glu Ala Asn Val Asp Phe Glu Ile Arg Phe Met Val
245 250 255

Asp Thr Asp Ile Val Gly Cys Asn Trp Leu Glu Leu Pro Ala Gly Lys
260 265 270

Tyr Ala Leu Arg Leu Lys Glu Lys Ala Thr Gln Cys Gln Leu Glu Ala
275 280 285

Asp Val Leu Trp Ser Asp Val Val Ser His Pro Pro Glu Gly Pro Trp
290 295 300

Gln Arg Ile Ala Pro Leu Arg Val Leu Ser Phe Asp Ile Glu Cys Ala
305 310 315 320

Gly Arg Lys Gly Ile Phe Pro Glu Pro Glu Arg Asp Pro Val Ile Gln
325 330 335

Ile Cys Ser Leu Gly Leu Arg Trp Gly Glu Pro Glu Pro Phe Leu Arg
340 345 350

Leu Ala Leu Thr Leu Arg Pro Cys Ala Pro Ile Leu Gly Ala Lys Val

355 360 365
 Gln Ser Tyr Glu Lys Glu Glu Asp Leu Leu Gln Ala Trp Ser Thr Phe
 370 375 380
 Ile Arg Ile Met Asp Pro Asp Val Ile Thr Gly Tyr Asn Ile Gln Asn
 385 390 395 400
 Phe Asp Leu Pro Tyr Leu Ile Ser Arg Ala Gln Thr Leu Lys Val Gln
 405 410 415
 Thr Phe Pro Phe Leu Gly Arg Val Ala Gly Leu Cys Ser Asn Ile Arg
 420 425 430
 Asp Ser Ser Phe Gln Ser Lys Gln Thr Gly Arg Arg Asp Thr Lys Val
 435 440 445
 Val Ser Met Val Gly Arg Val Gln Met Asp Met Leu Gln Val Leu Leu
 450 455 460
 Arg Glu Tyr Lys Leu Arg Ser His Thr Leu Asn Ala Val Ser Phe His
 465 470 475 480
 Phe Leu Gly Glu Gln Lys Glu Asp Val Gln His Ser Ile Ile Thr Asp
 485 490 495
 Leu Gln Asn Gly Asn Asp Gln Thr Arg Arg Arg Leu Ala Val Tyr Cys
 500 505 510
 Leu Lys Asp Ala Tyr Leu Pro Leu Arg Leu Leu Glu Arg Leu Met Val
 515 520 525
 Leu Val Asn Ala Val Glu Met Ala Arg Val Thr Gly Val Pro Leu Ser
 530 535 540
 Tyr Leu Leu Ser Arg Gly Gln Gln Val Lys Val Val Ser Gln Leu Leu
 545 550 555 560
 Arg Gln Ala Met His Glu Gly Leu Leu Met Pro Val Val Lys Ser Glu
 565 570 575
 Gly Gly Glu Asp Tyr Thr Gly Ala Thr Val Ile Glu Pro Leu Lys Gly
 580 585 590
 Tyr Tyr Asp Val Pro Ile Ala Thr Leu Asp Phe Ser Ser Leu Tyr Pro
 595 600 605
 Ser Ile Met Met Ala His Asn Leu Cys Tyr Thr Thr Leu Leu Arg Pro
 610 615 620

Gly Thr Ala Gln Lys Leu Gly Leu Thr Glu Asp Gln Phe Ile Arg Thr
625 630 635 640

Pro Thr Gly Asp Glu Phe Val Lys Thr Ser Val Arg Lys Gly Leu Leu
645 650 655

Pro Gln Ile Leu Glu Asn Leu Leu Ser Ala Arg Lys Arg Ala Lys Ala
660 665 670

Glu Leu Ala Lys Glu Thr Asp Pro Leu Arg Arg Gln Val Leu Asp Gly
675 680 685

Arg Gln Leu Ala Leu Lys Val Ser Ala Asn Ser Val Tyr Gly Phe Thr
690 695 700

Gly Ala Gln Val Gly Lys Leu Pro Cys Leu Glu Ile Ser Gln Ser Val
705 710 715 720

Thr Gly Phe Gly Arg Gln Met Ile Glu Lys Thr Lys Gln Leu Val Glu
725 730 735

Ser Lys Tyr Thr Val Glu Asn Gly Tyr Ser Thr Ser Ala Lys Val Val
740 745 750

Tyr Gly Asp Thr Asp Ser Val Met Cys Arg Phe Gly Val Ser Ser Val
755 760 765

Ala Glu Ala Met Ala Leu Gly Arg Glu Ala Ala Asp Trp Val Ser Gly
770 775 780

His Phe Pro Ser Pro Ile Arg Leu Glu Phe Glu Lys Val Tyr Phe Pro
785 790 795 800

Tyr Leu Leu Ile Ser Lys Lys Arg Tyr Ala Gly Leu Leu Phe Ser Ser
805 810 815

Arg Pro Asp Ala His Asp Arg Met Asp Cys Lys Gly Leu Glu Ala Val
820 825 830

Arg Arg Asp Asn Cys Pro Leu Val Ala Asn Leu Val Thr Ala Ser Leu
835 840 845

Arg Arg Leu Leu Ile Asp Arg Asp Pro Glu Gly Ala Val Ala His Ala
850 855 860

Gln Asp Val Ile Ser Asp Leu Leu Cys Asn Arg Ile Asp Ile Ser Gln
865 870 875 880

Leu Val Ile Thr Lys Glu Leu Thr Arg Ala Ala Ser Asp Tyr Ala Gly

	885	890	895
Lys Gln Ala His Val Glu Leu Ala Glu Arg Met Arg Lys Arg Asp Pro			
	900	905	910
Gly Ser Ala Pro Ser Leu Gly Asp Arg Val Pro Tyr Val Ile Ile Ser			
	915	920	925
Ala Ala Lys Gly Val Ala Ala Tyr Met Lys Ser Glu Asp Pro Leu Phe			
	930	935	940
Val Leu Glu His Ser Leu Pro Ile Asp Thr Gln Tyr Tyr Leu Glu Gln			
	945	950	955
Gln Leu Ala Lys Pro Leu Leu Arg Ile Phe Glu Pro Ile Leu Gly Glu			
	965	970	975
Gly Arg Ala Glu Ala Val Leu Leu Arg Gly Asp His Thr Arg Cys Lys			
	980	985	990
Thr Val Leu Thr Gly Lys Val Gly Gly Leu Leu Ala Phe Ala Lys Arg			
	995	1000	1005
Arg Asn Cys Cys Ile Gly Cys Arg Thr Val Leu Ser His Gln Gly Ala			
	1010	1015	1020
Val Cys Glu Phe Cys Gln Pro Arg Glu Ser Glu Leu Tyr Gln Lys Glu			
	1025	1030	1035
Val Ser His Leu Asn Ala Leu Glu Glu Arg Phe Ser Arg Leu Trp Thr			
	1045	1050	1055
Gln Cys Gln Arg Cys Gln Gly Ser Leu His Glu Asp Val Ile Cys Thr			
	1060	1065	1070
Ser Arg Asp Cys Pro Ile Phe Tyr Met Arg Lys Lys Val Arg Lys Asp			
	1075	1080	1085
Leu Glu Asp Gln Glu Gln Leu Leu Arg Arg Phe Gly Pro Pro Gly Pro			
	1090	1095	1100
Glu Ala Trp			
	1105		

(2) ANGABEN ZU SEQ ID NO: 23:

(i) SEQUENZKENNZEICHEN:

(A) LÄNGE: 781 Aminosäuren

(B) ART: Aminosäure

(C) STRANGFORM: Einzelstrang

(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

(A) ORGANISMUS: *Archaeoglobus fulgidus*

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 23:

Met Glu Arg Val Glu Gly Trp Leu Ile Asp Ala Asp Tyr Glu Thr Ile
 1 5 10 15

Gly Gly Lys Ala Val Val Arg Leu Trp Cys Lys Asp Asp Gln Gly Ile
 20 25 30

Phe Val Ala Tyr Asp Tyr Asn Phe Asp Pro Tyr Phe Tyr Val Ile Gly
 35 40 45

Val Asp Glu Asp Ile Leu Lys Asn Ala Ala Thr Ser Thr Arg Arg Glu
 50 55 60

Val Ile Lys Leu Lys Ser Phe Glu Lys Ala Gln Leu Lys Thr Leu Gly
 65 70 75 80

Arg Glu Val Glu Gly Tyr Ile Val Tyr Ala His His Pro Gln His Val
 85 90 95

Pro Lys Leu Arg Asp Tyr Leu Ser Gln Phe Gly Asp Val Arg Glu Ala
 100 105 110

Asp Ile Pro Phe Ala Tyr Arg Tyr Leu Ile Asp Lys Asp Leu Ala Cys
 115 120 125

Met Asp Gly Ile Ala Ile Glu Gly Glu Lys Gln Gly Gly Val Ile Arg
 130 135 140

Ser Tyr Lys Ile Glu Lys Val Glu Arg Ile Pro Arg Met Glu Phe Pro
 145 150 155 160

Glu Leu Lys Met Leu Val Phe Asp Cys Glu Met Leu Ser Ser Phe Gly
 165 170 175

Met Pro Glu Pro Glu Lys Asp Pro Ile Ile Val Ile Ser Val Lys Thr
 180 185 190

Asn Asp Asp Asp Glu Ile Ile Leu Thr Gly Asp Glu Arg Lys Ile Ile

195	200	205
Ser Asp Phe Val Lys Leu Ile Lys Ser Tyr Asp Pro Asp Ile Ile Val		
210	215	220
Gly Tyr Asn Gln Asp Ala Phe Asp Trp Pro Tyr Leu Arg Lys Arg Ala		
225	230	235 240
Glu Arg Trp Asn Ile Pro Leu Asp Val Gly Arg Asp Gly Ser Asn Val		
245	250	255
Val Phe Arg Gly Gly Arg Pro Lys Ile Thr Gly Arg Leu Asn Val Asp		
260	265	270
Leu Tyr Asp Ile Ala Met Arg Ile Ser Asp Ile Lys Ile Lys Lys Leu		
275	280	285
Glu Asn Val Ala Glu Phe Leu Gly Thr Lys Ile Glu Ile Ala Asp Ile		
290	295	300
Glu Ala Lys Asp Ile Tyr Arg Tyr Trp Ser Arg Gly Glu Lys Glu Lys		
305	310	315 320
Val Leu Asn Tyr Ala Arg Gln Asp Ala Ile Asn Thr Tyr Leu Ile Ala		
325	330	335
Lys Glu Leu Leu Pro Met His Tyr Glu Leu Ser Lys Met Ile Arg Leu		
340	345	350
Pro Val Asp Asp Val Thr Arg Met Gly Arg Gly Lys Gln Val Asp Trp		
355	360	365
Leu Leu Leu Ser Glu Ala Lys Lys Ile Gly Glu Ile Ala Pro Asn Pro		
370	375	380
Pro Glu His Ala Glu Ser Tyr Glu Gly Ala Phe Val Leu Glu Pro Glu		
385	390	395 400
Arg Gly Leu His Glu Asn Val Ala Cys Leu Asp Phe Ala Ser Met Tyr		
405	410	415
Pro Ser Ile Met Ile Ala Phe Asn Ile Ser Pro Asp Thr Tyr Gly Cys		
420	425	430
Arg Asp Asp Cys Tyr Glu Ala Pro Glu Val Gly His Lys Phe Arg Lys		
435	440	445
Ser Pro Asp Gly Phe Phe Lys Arg Ile Leu Arg Met Leu Ile Glu Lys		
450	455	460

Arg Arg Glu Leu Lys Val Glu Leu Lys Asn Leu Ser Pro Glu Ser Ser
 465 470 475 480

Glu Tyr Lys Leu Leu Asp Ile Lys Gln Gln Thr Leu Lys Val Leu Thr
 485 490 495

Asn Ser Phe Tyr Gly Tyr Met Gly Trp Asn Leu Ala Arg Trp Tyr Cys
 500 505 510

His Pro Cys Ala Glu Ala Thr Thr Ala Trp Gly Arg His Phe Ile Arg
 515 520 525

Thr Ser Ala Lys Ile Ala Glu Ser Met Gly Phe Lys Val Leu Tyr Gly
 530 535 540

Asp Thr Asp Ser Ile Phe Val Thr Lys Ala Gly Met Thr Lys Glu Asp
 545 550 555 560

Val Asp Arg Leu Ile Asp Lys Leu His Glu Glu Leu Pro Ile Gln Ile
 565 570 575

Glu Val Asp Glu Tyr Tyr Ser Ala Ile Phe Phe Val Glu Lys Lys Arg
 580 585 590

Tyr Ala Gly Leu Thr Glu Asp Gly Arg Leu Val Val Lys Gly Leu Glu
 595 600 605

Val Arg Arg Gly Asp Trp Cys Glu Leu Ala Lys Lys Val Gln Arg Glu
 610 615 620

Val Ile Glu Val Ile Leu Lys Glu Lys Asn Pro Glu Lys Ala Leu Ser
 625 630 635 640

Leu Val Lys Asp Val Ile Leu Arg Ile Lys Glu Gly Lys Val Ser Leu
 645 650 655

Glu Glu Val Val Ile Tyr Lys Gly Leu Thr Lys Lys Pro Ser Lys Tyr
 660 665 670

Glu Ser Met Gln Ala His Val Lys Ala Ala Leu Lys Ala Arg Glu Met
 675 680 685

Gly Ile Ile Tyr Pro Val Ser Ser Lys Ile Gly Tyr Val Ile Val Lys
 690 695 700

Gly Ser Gly Asn Ile Gly Asp Arg Ala Tyr Pro Ile Asp Leu Ile Glu
 705 710 715 720

Asp Phe Asp Gly Glu Asn Leu Arg Ile Lys Thr Lys Ser Gly Ile Glu

	725		730		735	
Ile	Lys	Lys	Leu	Asp	Lys	Asp
	740		745		750	
Tyr	Tyr	Ile	Asp	Asn	Gln	Ile
						Pro
Ser	Val	Leu	Arg	Ile	Leu	Glu
	755		760		765	
Arg	Phe	Gly	Tyr	Thr	Glu	Ala
						Ser
Leu						
Lys	Gly	Ser	Ser	Gln	Met	Ser
	770		775		780	
Leu	Asp	Ser	Phe	Phe	Ser	

(2) ANGABEN ZU SEQ ID NO: 24:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 1634 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: Methanococcus jannaschii

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 24:

Met	Gly	Met	Ser	Met	Gly	Lys	Ile	Lys	Ile	Asp	Ala	Leu	Ile	Asp	Asn
1		5		10		15									
Thr	Tyr	Lys	Thr	Ile	Glu	Asp	Lys	Ala	Val	Ile	Tyr	Leu	Tyr	Leu	Ile
	20		25		30										
Asn	Ser	Ile	Leu	Lys	Asp	Arg	Asp	Phe	Lys	Pro	Tyr	Phe	Tyr	Val	Glu
	35		40		45										
Leu	His	Lys	Glu	Lys	Val	Glu	Asn	Glu	Asp	Ile	Glu	Lys	Ile	Lys	Glu
	50		55		60										
Phe	Leu	Leu	Lys	Asn	Asp	Leu	Leu	Lys	Phe	Val	Glu	Asn	Ile	Glu	Val
	65		70		75		80								
Val	Lys	Lys	Ile	Ile	Leu	Arg	Lys	Glu	Lys	Glu	Val	Ile	Lys	Ile	Ile
	85		90		95										
Ala	Thr	His	Pro	Gln	Lys	Val	Pro	Lys	Leu	Arg	Lys	Ile	Lys	Glu	Cys
	100		105		110										

Glu Ile Val Lys Glu Ile Tyr Glu His Asp Ile Pro Phe Ala Lys Arg
 115 120 125

Tyr Leu Ile Asp Asn Glu Ile Ile Pro Met Thr Tyr Trp Asp Phe Glu
 130 135 140

Asn Lys Lys Pro Val Ser Ile Glu Ile Pro Lys Leu Lys Ser Val Ala
 145 150 155 160

Phe Asp Met Glu Val Tyr Asn Arg Asp Thr Glu Pro Asn Pro Glu Arg
 165 170 175

Asp Pro Ile Leu Met Ala Ser Phe Trp Asp Glu Asn Gly Gly Lys Val
 180 185 190

Ile Thr Tyr Lys Glu Phe Asn His Pro Asn Ile Glu Val Val Lys Asn
 195 200 205

Glu Lys Glu Leu Ile Lys Lys Ile Ile Glu Thr Leu Lys Glu Tyr Asp
 210 215 220

Val Ile Tyr Thr Tyr Asn Gly Asp Asn Phe Asp Phe Pro Tyr Leu Lys
 225 230 235 240

Ala Arg Ala Lys Ile Tyr Gly Ile Asp Ile Asn Leu Gly Lys Asp Gly
 245 250 255

Glu Glu Leu Lys Ile Lys Arg Gly Gly Met Glu Tyr Arg Ser Tyr Ile
 260 265 270

Pro Gly Arg Val His Ile Asp Leu Tyr Pro Ile Ser Arg Arg Leu Leu
 275 280 285

Lys Leu Thr Lys Tyr Thr Leu Glu Asp Val Val Tyr Asn Leu Phe Gly
 290 295 300

Ile Glu Lys Leu Lys Ile Pro His Thr Lys Ile Val Asp Tyr Trp Ala
 305 310 315 320

Asn Asn Asp Lys Thr Leu Ile Glu Tyr Ser Leu Gln Asp Ala Lys Tyr
 325 330 335

Thr Tyr Lys Ile Gly Lys Tyr Phe Phe Pro Leu Glu Val Met Phe Ser
 340 345 350

Arg Ile Val Asn Gln Thr Pro Phe Glu Ile Thr Arg Met Ser Ser Gly
 355 360 365

Gln Met Val Glu Tyr Leu Leu Met Lys Arg Ala Phe Lys Glu Asn Met

370 375 380
 Ile Val Pro Asn Lys Pro Asp Glu Glu Glu Tyr Arg Arg Arg Val Leu
 385 390 395 400
 Thr Thr Tyr Glu Gly Gly Tyr Val Lys Glu Pro Glu Lys Gly Met Phe
 405 410 415
 Glu Asp Ile Ile Ser Met Asp Phe Arg Cys His Pro Lys Gly Thr Lys
 420 425 430
 Val Val Val Lys Gly Lys Gly Ile Val Asn Ile Glu Asp Val Lys Glu
 435 440 445
 Gly Asn Tyr Val Leu Gly Ile Asp Gly Trp Gln Lys Val Lys Lys Val
 450 455 460
 Trp Lys Tyr Glu Tyr Glu Gly Glu Leu Ile Asn Val Asn Gly Leu Lys
 465 470 475 480
 Cys Thr Pro Asn His Lys Ile Pro Leu Arg Tyr Lys Ile Lys His Lys
 485 490 495
 Lys Ile Asn Lys Asn Asp Tyr Leu Val Arg Asp Ile Tyr Ala Lys Ser
 500 505 510
 Leu Leu Thr Lys Phe Lys Gly Glu Gly Lys Leu Ile Leu Cys Lys Asp
 515 520 525
 Phe Glu Thr Ile Gly Asn Tyr Glu Lys Tyr Ile Asn Asp Met Asp Glu
 530 535 540
 Asp Phe Ile Leu Lys Ser Glu Leu Ile Gly Ile Leu Leu Ala Glu Gly
 545 550 555 560
 His Leu Leu Arg Arg Asp Ile Glu Tyr Phe Asp Ser Ser Arg Gly Lys
 565 570 575
 Lys Arg Ile Ser His Gln Tyr Arg Val Glu Ile Thr Val Asn Glu Asp
 580 585 590
 Glu Lys Asp Phe Ile Glu Lys Ile Lys Tyr Ile Phe Lys Lys Leu Phe
 595 600 605
 Asn Tyr Glu Leu Tyr Val Arg Arg Lys Lys Gly Thr Lys Ala Ile Thr
 610 615 620
 Leu Gly Cys Ala Lys Lys Asp Ile Tyr Leu Lys Ile Glu Glu Ile Leu
 625 630 635 640

Lys Asn Lys Glu Lys Tyr Leu Pro Asn Ala Ile Leu Arg Gly Phe Phe
 645 650 655

Glu Gly Asp Gly Tyr Val Asn Thr Val Arg Arg Ala Val Val Val Asn
 660 665 670

Gln Gly Thr Asn Asn Tyr Asp Lys Ile Lys Phe Ile Ala Ser Leu Leu
 675 680 685

Asp Arg Leu Gly Ile Lys Tyr Ser Phe Tyr Thr Tyr Ser Tyr Glu Glu
 690 695 700

Arg Gly Lys Lys Leu Lys Arg Tyr Val Ile Glu Ile Phe Ser Lys Gly
 705 710 715 720

Asp Leu Ile Lys Phe Ser Ile Leu Ile Ser Phe Ile Ser Arg Arg Lys
 725 730 735

Asn Asn Leu Leu Asn Glu Ile Ile Arg Gln Lys Thr Leu Tyr Lys Ile
 740 745 750

Gly Asp Tyr Gly Phe Tyr Asp Leu Asp Asp Val Cys Val Ser Leu Glu
 755 760 765

Ser Tyr Lys Gly Glu Val Tyr Asp Leu Thr Leu Glu Gly Arg Pro Tyr
 770 775 780

Tyr Phe Ala Asn Gly Ile Leu Thr His Asn Ser Leu Tyr Pro Ser Ile
 785 790 795 800

Ile Ile Ser Tyr Asn Ile Ser Pro Asp Thr Leu Asp Cys Glu Cys Cys
 805 810 815

Lys Asp Val Ser Glu Lys Ile Leu Gly His Trp Phe Cys Lys Lys Lys
 820 825 830

Glu Gly Leu Ile Pro Lys Thr Leu Arg Asn Leu Ile Glu Arg Arg Ile
 835 840 845

Asn Ile Lys Arg Arg Met Lys Lys Met Ala Glu Ile Gly Glu Ile Asn
 850 855 860

Glu Glu Tyr Asn Leu Leu Asp Tyr Glu Gln Lys Ser Leu Lys Ile Leu
 865 870 875 880

Ala Asn Ser Ile Leu Pro Asp Glu Tyr Leu Thr Ile Ile Glu Glu Asp
 885 890 895

Gly Ile Lys Val Val Lys Ile Gly Glu Tyr Ile Asp Asp Leu Met Arg

900	905	910
Lys His Lys Asp Lys Ile Lys Phe Ser Gly Ile Ser Glu Ile Leu Glu		
915	920	925
Thr Lys Asn Leu Lys Thr Phe Ser Phe Asp Lys Ile Thr Lys Lys Cys		
930	935	940
Glu Ile Lys Lys Val Lys Ala Leu Ile Arg His Pro Tyr Phe Gly Lys		
945	950	955
Ala Tyr Lys Ile Lys Leu Arg Ser Gly Arg Thr Ile Lys Val Thr Arg		
965	970	975
Gly His Ser Leu Phe Lys Tyr Glu Asn Gly Lys Ile Val Glu Val Lys		
980	985	990
Gly Asp Asp Val Arg Phe Gly Asp Leu Ile Val Val Pro Lys Lys Leu		
995	1000	1005
Thr Cys Val Asp Lys Glu Val Val Ile Asn Ile Pro Lys Arg Leu Ile		
1010	1015	1020
Asn Ala Asp Glu Glu Glu Ile Lys Asp Leu Val Ile Thr Lys His Lys		
1025	1030	1035
Asp Lys Ala Phe Phe Val Lys Leu Lys Lys Thr Leu Glu Asp Ile Glu		
1045	1050	1055
Asn Asn Lys Leu Lys Val Ile Phe Asp Asp Cys Ile Leu Tyr Leu Lys		
1060	1065	1070
Glu Leu Gly Leu Ile Asp Tyr Asn Ile Ile Lys Lys Ile Asn Lys Val		
1075	1080	1085
Asp Ile Lys Ile Leu Asp Glu Glu Lys Phe Lys Ala Tyr Lys Lys Tyr		
1090	1095	1100
Phe Asp Thr Val Ile Glu His Gly Asn Phe Lys Lys Gly Arg Cys Asn		
1105	1110	1115
Ile Gln Tyr Ile Lys Ile Lys Asp Tyr Ile Ala Asn Ile Pro Asp Lys		
1125	1130	1135
Glu Phe Glu Asp Cys Glu Ile Gly Ala Tyr Ser Gly Lys Ile Asn Ala		
1140	1145	1150
Leu Leu Lys Leu Asp Glu Lys Leu Ala Lys Phe Leu Gly Phe Phe Val		
1155	1160	1165

Thr Arg Gly Arg Leu Lys Lys Gln Lys Leu Lys Gly Glu Thr Val Tyr
 1170 1175 1180

Glu Ile Ser Val Tyr Lys Ser Leu Pro Glu Tyr Gln Lys Glu Ile Ala
 1185 1190 1195 1200

Glu Thr Phe Lys Glu Val Phe Gly Ala Gly Ser Met Val Lys Asp Lys
 1205 1210 1215

Val Thr Met Asp Asn Lys Ile Val Tyr Leu Val Leu Lys Tyr Ile Phe
 1220 1225 1230

Lys Cys Gly Asp Lys Asp Lys Lys His Ile Pro Glu Glu Leu Phe Leu
 1235 1240 1245

Ala Ser Glu Ser Val Ile Lys Ser Phe Leu Asp Gly Phe Leu Lys Ala
 1250 1255 1260

Lys Lys Asn Ser His Lys Gly Thr Ser Thr Phe Met Ala Lys Asp Glu
 1265 1270 1275 1280

Lys Tyr Leu Asn Gln Leu Met Ile Leu Phe Asn Leu Val Gly Ile Pro
 1285 1290 1295

Thr Arg Phe Thr Pro Val Lys Asn Lys Gly Tyr Lys Leu Thr Leu Asn
 1300 1305 1310

Pro Lys Tyr Gly Thr Val Lys Asp Leu Met Leu Asp Glu Val Lys Glu
 1315 1320 1325

Ile Glu Ala Phe Glu Tyr Ser Gly Tyr Val Tyr Asp Leu Ser Val Glu
 1330 1335 1340

Asp Asn Glu Asn Phe Leu Val Asn Asn Ile Tyr Ala His Asn Ser Val
 1345 1350 1355 1360

Tyr Gly Tyr Leu Ala Phe Pro Arg Ala Arg Phe Tyr Ser Arg Glu Cys
 1365 1370 1375

Ala Glu Ile Val Thr Tyr Leu Gly Arg Lys Tyr Ile Leu Glu Thr Val
 1380 1385 1390

Lys Glu Ala Glu Lys Phe Gly Phe Lys Val Leu Tyr Ile Asp Thr Asp
 1395 1400 1405

Gly Phe Tyr Ala Ile Trp Lys Glu Lys Ile Ser Lys Glu Glu Leu Ile
 1410 1415 1420

Lys Lys Ala Met Glu Phe Val Glu Tyr Ile Asn Ser Lys Leu Pro Gly

1425 1430 1435 1440
 Thr Met Glu Leu Glu Phe Glu Gly Tyr Phe Lys Arg Gly Ile Phe Val
 1445 1450 1455
 Thr Lys Lys Arg Tyr Ala Leu Ile Asp Glu Asn Gly Arg Val Thr Val
 1460 1465 1470
 Lys Gly Leu Glu Phe Val Arg Arg Asp Trp Ser Asn Ile Ala Lys Ile
 1475 1480 1485
 Thr Gln Arg Arg Val Leu Glu Ala Leu Leu Val Glu Gly Ser Ile Glu
 1490 1495 1500
 Lys Ala Lys Lys Ile Ile Gln Asp Val Ile Lys Asp Leu Arg Glu Lys
 1505 1510 1515 1520
 Lys Ile Lys Lys Glu Asp Leu Ile Ile Tyr Thr Gln Leu Thr Lys Asp
 1525 1530 1535
 Pro Lys Glu Tyr Lys Thr Thr Ala Pro His Val Glu Ile Ala Lys Lys
 1540 1545 1550
 Leu Met Arg Glu Gly Lys Arg Ile Lys Val Gly Asp Ile Ile Gly Tyr
 1555 1560 1565
 Ile Ile Val Lys Gly Thr Lys Ser Ile Ser Glu Arg Ala Lys Leu Pro
 1570 1575 1580
 Glu Glu Val Asp Ile Asp Asp Ile Asp Val Asn Tyr Tyr Ile Asp Asn
 1585 1590 1595 1600
 Gln Ile Leu Pro Pro Val Leu Arg Ile Met Glu Ala Val Gly Val Ser
 1605 1610 1615
 Lys Asn Glu Leu Lys Lys Glu Gly Ala Gln Leu Thr Leu Asp Lys Phe
 1620 1625 1630
 Phe Lys

(2) ANGABEN ZU SEQ ID NO: 25:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 1235 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

(A) ORGANISMUS: *Pyrococcus horikoshii*

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 25:

Met Ile Leu Asp Ala Asp Tyr Ile Thr Glu Asp Gly Lys Pro Ile Ile
1 5 10 15

Arg Ile Phe Lys Lys Glu Asn Gly Glu Phe Lys Val Glu Tyr Asp Arg
20 25 30

Asn Phe Arg Pro Tyr Ile Tyr Ala Leu Leu Arg Asp Asp Ser Ala Ile
35 40 45

Asp Glu Ile Lys Lys Ile Thr Ala Gln Arg His Gly Lys Val Val Arg
50 55 60

Ile Val Glu Thr Glu Lys Ile Gln Arg Lys Phe Leu Gly Arg Pro Ile
65 70 75 80

Glu Val Trp Lys Leu Tyr Leu Glu His Pro Gln Asp Val Pro Ala Ile
85 90 95

Arg Asp Lys Ile Arg Glu His Pro Ala Val Val Asp Ile Phe Glu Tyr
100 105 110

Asp Ile Pro Phe Ala Lys Arg Tyr Leu Ile Asp Lys Gly Leu Thr Pro
115 120 125

Met Glu Gly Asn Glu Lys Leu Thr Phe Leu Ala Val Asp Ile Glu Thr
130 135 140

Leu Tyr His Glu Gly Glu Glu Phe Gly Lys Gly Pro Val Ile Met Ile
145 150 155 160

Ser Tyr Ala Asp Glu Glu Gly Ala Lys Val Ile Thr Trp Lys Lys Ile
165 170 175

Asp Leu Pro Tyr Val Glu Val Val Ser Ser Glu Arg Glu Met Ile Lys
180 185 190

Arg Leu Ile Arg Val Ile Lys Glu Lys Asp Pro Asp Val Ile Ile Thr
195 200 205

Tyr Asn Gly Asp Asn Phe Asp Phe Pro Tyr Leu Leu Lys Arg Ala Glu

210 215 220
 Lys Leu Gly Ile Lys Leu Leu Leu Gly Arg Asp Asn Ser Glu Pro Lys
 225 230 235 240
 Met Gln Lys Met Gly Asp Ser Leu Ala Val Glu Ile Lys Gly Arg Ile
 245 250 255
 His Phe Asp Leu Phe Pro Val Ile Arg Arg Thr Ile Asn Leu Pro Thr
 260 265 270
 Tyr Thr Leu Glu Ala Val Tyr Glu Ala Ile Phe Gly Lys Pro Lys Glu
 275 280 285
 Lys Val Tyr Ala Asp Glu Ile Ala Lys Ala Trp Glu Thr Gly Glu Gly
 290 295 300
 Leu Glu Arg Val Ala Lys Tyr Ser Met Glu Asp Ala Lys Val Thr Tyr
 305 310 315 320
 Glu Leu Gly Arg Glu Phe Phe Pro Met Glu Ala Gln Leu Ala Arg Leu
 325 330 335
 Val Gly Gln Pro Val Trp Asp Val Ser Arg Ser Ser Thr Gly Asn Leu
 340 345 350
 Val Glu Trp Phe Leu Leu Arg Lys Ala Tyr Glu Arg Asn Glu Leu Ala
 355 360 365
 Pro Asn Lys Pro Asp Glu Lys Glu Tyr Glu Arg Arg Leu Arg Glu Ser
 370 375 380
 Tyr Glu Gly Gly Tyr Val Lys Glu Pro Glu Lys Gly Leu Trp Glu Gly
 385 390 395 400
 Ile Val Ser Leu Asp Phe Arg Ser Leu Tyr Pro Ser Ile Ile Ile Thr
 405 410 415
 His Asn Val Ser Pro Asp Thr Leu Asn Arg Glu Gly Cys Glu Glu Tyr
 420 425 430
 Asp Val Ala Pro Lys Val Gly His Arg Phe Cys Lys Asp Phe Pro Gly
 435 440 445
 Phe Ile Pro Ser Leu Leu Gly Gln Leu Leu Glu Glu Arg Gln Lys Ile
 450 455 460
 Lys Lys Arg Met Lys Glu Ser Lys Asp Pro Val Glu Lys Lys Leu Leu
 465 470 475 480

Asp Tyr Arg Gln Arg Ala Ile Lys Ile Leu Ala Asn Ser Ile Leu Pro
 485 490 495

Asp Glu Trp Leu Pro Ile Val Glu Asn Glu Lys Val Arg Phe Val Lys
 500 505 510

Ile Gly Asp Phe Ile Asp Arg Glu Ile Glu Glu Asn Ala Glu Arg Val
 515 520 525

Lys Arg Asp Gly Glu Thr Glu Ile Leu Glu Val Lys Asp Leu Lys Ala
 530 535 540

Leu Ser Phe Asn Arg Glu Thr Lys Lys Ser Glu Leu Lys Lys Val Lys
 545 550 555 560

Ala Leu Ile Arg His Arg Tyr Ser Gly Lys Val Tyr Ser Ile Lys Leu
 565 570 575

Lys Ser Gly Arg Arg Ile Lys Ile Thr Ser Gly His Ser Leu Phe Ser
 580 585 590

Val Lys Asn Gly Lys Leu Val Lys Val Arg Gly Asp Glu Leu Lys Pro
 595 600 605

Gly Asp Leu Val Val Val Pro Gly Arg Leu Lys Leu Pro Glu Ser Lys
 610 615 620

Gln Val Leu Asn Leu Val Glu Leu Leu Lys Leu Pro Glu Glu Glu
 625 630 635 640

Thr Ser Asn Ile Val Met Met Ile Pro Val Lys Gly Arg Lys Asn Phe
 645 650 655

Phe Lys Gly Met Leu Lys Thr Leu Tyr Trp Ile Phe Gly Glu Gly Glu
 660 665 670

Arg Pro Arg Thr Ala Gly Arg Tyr Leu Lys His Leu Glu Arg Leu Gly
 675 680 685

Tyr Val Lys Leu Lys Arg Arg Gly Cys Glu Val Leu Asp Trp Glu Ser
 690 695 700

Leu Lys Arg Tyr Arg Lys Leu Tyr Glu Thr Leu Ile Lys Asn Leu Lys
 705 710 715 720

Tyr Asn Gly Asn Ser Arg Ala Tyr Met Val Glu Phe Asn Ser Leu Arg
 725 730 735

Asp Val Val Ser Leu Met Pro Ile Glu Glu Leu Lys Glu Trp Ile Ile

740 745 750
 Gly Glu Pro Arg Gly Pro Lys Ile Gly Thr Phe Ile Asp Val Asp Asp
 755 760 765
 Ser Phe Ala Lys Leu Leu Gly Tyr Tyr Ile Ser Ser Gly Asp Val Glu
 770 775 780
 Lys Asp Arg Val Lys Phe His Ser Lys Asp Gln Asn Val Leu Glu Asp
 785 790 795 800
 Ile Ala Lys Leu Ala Glu Lys Leu Phe Gly Lys Val Arg Arg Gly Arg
 805 810 815
 Gly Tyr Ile Glu Val Ser Gly Lys Ile Ser His Ala Ile Phe Arg Val
 820 825 830
 Leu Ala Glu Gly Lys Arg Ile Pro Glu Phe Ile Phe Thr Ser Pro Met
 835 840 845
 Asp Ile Lys Val Ala Phe Leu Lys Gly Leu Asn Gly Asn Ala Glu Glu
 850 855 860
 Leu Thr Phe Ser Thr Lys Ser Glu Leu Leu Val Asn Gln Leu Ile Leu
 865 870 875 880
 Leu Leu Asn Ser Ile Gly Val Ser Asp Ile Lys Ile Glu His Glu Lys
 885 890 895
 Gly Val Tyr Arg Val Tyr Ile Asn Lys Lys Glu Ser Ser Asn Gly Asp
 900 905 910
 Ile Val Leu Asp Ser Val Glu Ser Ile Glu Val Glu Lys Tyr Glu Gly
 915 920 925
 Tyr Val Tyr Asp Leu Ser Val Glu Asp Asn Glu Asn Phe Leu Val Gly
 930 935 940
 Phe Gly Leu Leu Tyr Ala His Asn Ser Tyr Tyr Gly Tyr Tyr Gly Tyr
 945 950 955 960
 Ala Lys Ala Arg Trp Tyr Cys Lys Glu Cys Ala Glu Ser Val Thr Ala
 965 970 975
 Trp Gly Arg Gln Tyr Ile Asp Leu Val Arg Arg Glu Leu Glu Ala Arg
 980 985 990
 Gly Phe Lys Val Leu Tyr Ile Asp Thr Asp Gly Leu Tyr Ala Thr Ile
 995 1000 1005

Pro Gly Val Lys Asp Trp Glu Glu Val Lys Arg Arg Ala Leu Glu Phe
 1010 1015 1020

Val Asp Tyr Ile Asn Ser Lys Leu Pro Gly Val Leu Glu Leu Glu Tyr
 1025 1030 1035 1040

Glu Gly Phe Tyr Ala Arg Gly Phe Phe Val Thr Lys Lys Lys Tyr Ala
 1045 1050 1055

Leu Ile Asp Glu Glu Gly Lys Ile Val Thr Arg Gly Leu Glu Ile Val
 1060 1065 1070

Arg Arg Asp Trp Ser Glu Ile Ala Lys Glu Thr Gln Ala Arg Val Leu
 1075 1080 1085

Glu Ala Ile Leu Lys His Gly Asn Val Glu Glu Ala Val Lys Ile Val
 1090 1095 1100

Lys Asp Val Thr Glu Lys Leu Thr Asn Tyr Glu Val Pro Pro Glu Lys
 1105 1110 1115 1120

Leu Val Ile Tyr Glu Gln Ile Thr Arg Pro Ile Asn Glu Tyr Lys Ala
 1125 1130 1135

Ile Gly Pro His Val Ala Val Ala Lys Arg Leu Met Ala Arg Gly Ile
 1140 1145 1150

Lys Val Lys Pro Gly Met Val Ile Gly Tyr Ile Val Leu Arg Gly Asp
 1155 1160 1165

Gly Pro Ile Ser Lys Arg Ala Ile Ser Ile Glu Glu Phe Asp Pro Arg
 1170 1175 1180

Lys His Lys Tyr Asp Ala Glu Tyr Tyr Ile Glu Asn Gln Val Leu Pro
 1185 1190 1195 1200

Ala Val Glu Arg Ile Leu Lys Ala Phe Gly Tyr Lys Arg Glu Asp Leu
 1205 1210 1215

Arg Trp Gln Lys Thr Lys Gln Val Gly Leu Gly Ala Trp Ile Lys Val
 1220 1225 1230

Lys Lys Ser
 1235

(2) ANGABEN ZU SEQ ID NO: 26:

(i) SEQUENZKENNZEICHEN:

(A) LÄNGE: 586 Aminosäuren

- (B) ART: Aminosäure
(C) STRANGFORM: Einzelstrang
(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

(A) ORGANISMUS: Methanobacterium thermoautotrophicum

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 26:

Met Glu Asp Tyr Arg Met Val Leu Leu Asp Ile Asp Tyr Val Thr Val
1 5 10 15

Asp Glu Val Pro Val Ile Arg Leu Phe Gly Lys Asp Lys Ser Gly Gly
20 25 30

Asn Glu Pro Ile Ile Ala His Asp Arg Ser Phe Arg Pro Tyr Ile Tyr
35 40 45

Ala Ile Pro Thr Asp Leu Asp Glu Cys Leu Arg Glu Leu Glu Glu Leu
50 55 60

Glu Leu Glu Lys Leu Glu Val Lys Glu Met Arg Asp Leu Gly Arg Pro
65 70 75 80

Thr Glu Val Ile Arg Ile Glu Phe Arg His Pro Gln Asp Val Pro Lys
85 90 95

Ile Arg Asp Arg Ile Arg Asp Leu Glu Ser Val Arg Asp Ile Arg Glu
100 105 110

His Asp Ile Pro Phe Tyr Arg Arg Tyr Leu Ile Asp Lys Ser Ile Val
115 120 125

Pro Met Glu Glu Leu Glu Phe Gln Gly Val Glu Val Asp Ser Ala Pro
130 135 140

Ser Val Thr Thr Asp Val Arg Thr Val Glu Val Thr Gly Arg Val Gln
145 150 155 160

Ser Thr Gly Ser Gly Ala His Gly Leu Asp Ile Leu Ser Phe Asp Ile
165 170 175

Glu Val Arg Asn Pro His Gly Met Pro Asp Pro Glu Lys Asp Glu Ile
180 185 190

Val Met Ile Gly Val Ala Gly Asn Met Gly Tyr Glu Ser Val Ile Ser
195 200 205

Thr Ala Gly Asp His Leu Asp Phe Val Glu Val Val Glu Asp Glu Arg
210 215 220

Glu Leu Leu Glu Arg Phe Ala Glu Ile Val Ile Asp Lys Lys Pro Asp
225 230 235 240

Ile Leu Val Gly Tyr Asn Ser Asp Asn Phe Asp Phe Pro Tyr Ile Thr
245 250 255

Arg Arg Ala Ala Ile Leu Gly Ala Glu Leu Asp Leu Gly Trp Asp Gly
260 265 270

Ser Lys Ile Arg Thr Met Arg Arg Gly Phe Ala Asn Ala Thr Ala Ile
275 280 285

Lys Gly Thr Val His Val Asp Leu Tyr Pro Val Met Arg Arg Tyr Met
290 295 300

Asn Leu Asp Arg Tyr Thr Leu Glu Arg Val Tyr Gln Glu Leu Phe Gly
305 310 315 320

Glu Glu Lys Ile Asp Leu Pro Gly Asp Arg Leu Trp Glu Tyr Trp Asp
325 330 335

Arg Asp Glu Leu Arg Asp Glu Leu Phe Arg Tyr Ser Leu Asp Asp Val
340 345 350

Val Ala Thr His Arg Ile Ala Glu Lys Ile Leu Pro Leu Asn Leu Glu
355 360 365

Leu Thr Arg Leu Val Gly Gln Pro Leu Phe Asp Ile Ser Arg Met Ala
370 375 380

Thr Gly Gln Gln Ala Glu Trp Phe Leu Val Arg Lys Ala Tyr Gln Tyr
385 390 395 400

Gly Glu Leu Val Pro Asn Lys Pro Ser Gln Ser Asp Phe Ser Ser Arg
405 410 415

Arg Gly Arg Arg Ala Val Gly Gly Tyr Val Lys Glu Pro Glu Lys Gly
420 425 430

Leu His Glu Asn Ile Val Gln Phe Asp Phe Arg Ser Leu Tyr Pro Ser
435 440 445

Ile Ile Ile Ser Lys Asn Ile Ser Pro Asp Thr Leu Thr Asp Asp Glu

450 455 460
 Glu Ser Glu Cys Tyr Val Ala Pro Glu Tyr Gly Tyr Arg Phe Arg Lys
 465 470 475 480
 Ser Pro Arg Gly Phe Val Pro Ser Val Ile Gly Glu Ile Leu Ser Glu
 485 490 495
 Arg Val Arg Ile Lys Glu Glu Met Lys Gly Ser Asp Asp Pro Met Glu
 500 505 510
 Arg Lys Ile Leu Asn Val Gln Gln Glu Ala Leu Lys Arg Leu Ala Asn
 515 520 525
 Thr Met Tyr Gly Val Tyr Gly Tyr Ser Arg Phe Arg Trp Tyr Ser Met
 530 535 540
 Glu Cys Ala Glu Ala Ile Thr Ala Trp Gly Arg Asp Tyr Ile Lys Lys
 545 550 555 560
 Thr Ile Lys Thr Ala Glu Glu Phe Gly Phe His Thr Val Tyr Ala Asp
 565 570 575
 Thr Asp Gly Phe Tyr Ala Thr Tyr Arg Gly
 580 585

(2) ANGABEN ZU SEQ ID NO: 27:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 1143 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: *Archaeoglobus fulgidus*

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 27:

Met Asp Ala Thr Leu Asp Arg Phe Phe Pro Leu Phe Glu Ser Glu Ser
 1 5 10 15
 Asn Glu Asp Phe Trp Arg Ile Glu Glu Ile Arg Arg Tyr His Glu Ser
 20 25 30

Leu Met Val Glu Leu Asp Arg Ile Tyr Arg Ile Ala Glu Ala Ala Arg
 35 40 45

Lys Lys Gly Leu Asp Pro Glu Leu Ser Val Glu Ile Pro Ile Ala Lys
 50 55 60

Asn Met Ala Glu Arg Val Glu Lys Leu Met Asn Leu Gln Gly Leu Ala
 65 70 75 80

Lys Arg Ile Met Glu Leu Glu Glu Gly Gly Leu Ser Arg Glu Leu Ile
 85 90 95

Cys Phe Lys Val Ala Asp Glu Ile Val Glu Gly Lys Phe Gly Glu Met
 100 105 110

Pro Lys Glu Glu Ala Ile Asp Lys Ala Val Arg Thr Ala Val Ala Ile
 115 120 125

Met Thr Glu Gly Val Val Ala Ala Pro Ile Glu Gly Ile Ala Arg Val
 130 135 140

Arg Ile Asp Arg Glu Asn Phe Leu Arg Val Tyr Tyr Ala Gly Pro Ile
 145 150 155 160

Arg Ser Ala Gly Gly Thr Ala Gln Val Ile Ser Val Leu Val Ala Asp
 165 170 175

Tyr Val Arg Arg Lys Ala Glu Ile Gly Arg Tyr Val Pro Thr Glu Glu
 180 185 190

Glu Ile Leu Arg Tyr Cys Glu Glu Ile Pro Leu Tyr Lys Lys Val Ala
 195 200 205

Asn Leu Gln Tyr Leu Pro Ser Asp Glu Glu Ile Arg Leu Ile Val Ser
 210 215 220

Asn Cys Pro Ile Cys Ile Asp Gly Glu Pro Thr Glu Ser Ala Glu Val
 225 230 235 240

Ser Gly Tyr Arg Asn Leu Pro Arg Val Glu Thr Asn Arg Val Arg Gly
 245 250 255

Gly Met Ala Leu Val Ile Ala Glu Gly Ile Ala Leu Lys Ala Pro Lys
 260 265 270

Leu Lys Lys Met Val Asp Glu Val Gly Ile Glu Gly Trp Glu Trp Leu
 275 280 285

Asp Ala Leu Ile Lys Gly Gly Gly Asp Ser Gly Ser Glu Glu Glu Lys

290	295	300	
Ala Val Ile Lys Pro Lys Asp Lys Tyr Leu Ser Asp Ile Val Ala Gly			
305	310	315	320
Arg Pro Val Leu Ser His Pro Ser Arg Lys Gly Gly Phe Arg Leu Arg			
	325	330	335
Tyr Gly Arg Ala Arg Asn Ser Gly Phe Ala Thr Val Gly Val Asn Pro			
	340	345	350
Ala Thr Met Tyr Leu Leu Glu Phe Val Ala Val Gly Thr Gln Leu Lys			
	355	360	365
Val Glu Arg Pro Gly Lys Ala Gly Gly Val Val Pro Val Ser Thr Ile			
	370	375	380
Glu Gly Pro Thr Val Arg Leu Lys Asn Gly Asp Val Val Lys Ile Asn			
385	390	395	400
Thr Leu Ser Glu Ala Lys Ala Leu Lys Gly Glu Val Ala Ala Ile Leu			
	405	410	415
Asp Leu Gly Glu Ile Leu Ile Asn Tyr Gly Asp Phe Leu Glu Asn Asn			
	420	425	430
His Pro Leu Ile Pro Ala Ser Tyr Thr Tyr Glu Trp Trp Ile Gln Glu			
	435	440	445
Ala Glu Lys Ala Gly Leu Arg Gly Asp Tyr Arg Lys Ile Ser Glu Glu			
	450	455	460
Glu Ala Leu Lys Leu Cys Asp Glu Phe His Val Pro Leu His Pro Asp			
465	470	475	480
Tyr Thr Tyr Leu Trp His Asp Ile Ser Val Glu Asp Tyr Arg Tyr Leu			
	485	490	495
Arg Asn Phe Val Ser Asp Asn Gly Lys Ile Glu Gly Lys His Gly Lys			
	500	505	510
Ser Val Leu Leu Leu Pro Tyr Asp Ser Arg Val Lys Glu Ile Leu Glu			
	515	520	525
Ala Leu Leu L u Glu His Lys Val Arg Glu Ser Phe Ile Val Ile Glu			
	530	535	540
Thr Trp Arg Ala Phe Ile Arg Cys Leu Gly Leu Asp Glu Lys Leu Ser			
545	550	555	560

Lys Val Ser Glu Val Ser Gly Lys Asp Val Leu Glu Ile Val Asn Gly
565 570 575

Ile Ser Gly Ile Lys Val Arg Pro Lys Ala Leu Ser Arg Ile Gly Ala
580 585 590

Arg Met Gly Arg Pro Glu Lys Ala Lys Glu Arg Lys Met Ser Pro Pro
595 600 605

Pro His Ile Leu Phe Pro Val Gly Met Ala Gly Gly Asn Thr Arg Asp
610 615 620

Ile Lys Asn Ala Ile Asn Tyr Thr Lys Ser Tyr Asn Ala Lys Lys Gly
625 630 635 640

Glu Ile Glu Val Glu Ile Ala Ile Arg Lys Cys Pro Gln Cys Gly Lys
645 650 655

Glu Thr Phe Trp Leu Lys Cys Asp Val Cys Gly Glu Leu Thr Glu Gln
660 665 670

Leu Tyr Tyr Cys Pro Ser Cys Arg Met Lys Asn Thr Ser Ser Val Cys
675 680 685

Glu Ser Cys Gly Arg Glu Cys Glu Gly Tyr Met Lys Arg Lys Val Asp
690 695 700

Leu Arg Glu Leu Tyr Glu Glu Ala Ile Ala Asn Leu Gly Glu Tyr Asp
705 710 715 720

Ser Phe Asp Thr Ile Lys Gly Val Lys Gly Met Thr Ser Lys Thr Lys
725 730 735

Ile Pro Glu Arg Leu Glu Lys Gly Ile Leu Arg Val Lys His Gly Val
740 745 750

Phe Val Phe Lys Asp Gly Thr Ala Arg Phe Asp Ala Thr Asp Leu Pro
755 760 765

Ile Thr His Phe Lys Pro Ala Glu Ile Gly Val Ser Val Glu Lys Leu
770 775 780

Arg Glu Leu Gly Tyr Glu Arg Asp Tyr Lys Gly Ala Glu Leu Lys Asn
785 790 795 800

Glu Asn Gln Ile Val Glu Leu Lys Pro Gln Asp Val Ile Leu Pro Lys
805 810 815

Ser Gly Ala Glu Tyr Leu Leu Arg Val Ala Asn Phe Ile Asp Asp Leu

820 825 830
 Leu Val Lys Phe Tyr Lys Met Glu Pro Phe Tyr Asn Ala Lys Ser Val
 835 840 845
 Glu Asp Leu Ile Gly His Leu Val Ile Gly Leu Ala Pro His Thr Ser
 850 855 860
 Ala Gly Val Leu Gly Arg Ile Ile Gly Phe Ser Asp Val Leu Ala Gly
 865 870 875 880
 Tyr Ala His Pro Tyr Phe His Ala Ala Lys Arg Arg Asn Cys Asp Gly
 885 890 895
 Asp Glu Asp Cys Phe Met Leu Leu Leu Asp Gly Leu Leu Asn Phe Ser
 900 905 910
 Arg Lys Phe Leu Pro Asp Lys Arg Gly Gly Gln Met Asp Ala Pro Leu
 915 920 925
 Val Leu Thr Ala Ile Val Asp Pro Arg Glu Val Asp Lys Glu Val His
 930 935 940
 Asn Met Asp Ile Val Glu Arg Tyr Pro Leu Glu Phe Tyr Glu Ala Thr
 945 950 955 960
 Met Arg Phe Ala Ser Pro Lys Glu Met Glu Asp Tyr Val Glu Lys Val
 965 970 975
 Lys Asp Arg Leu Lys Asp Glu Ser Arg Phe Cys Gly Leu Phe Phe Thr
 980 985 990
 His Asp Thr Glu Asn Ile Ala Ala Gly Val Lys Glu Ser Ala Tyr Lys
 995 1000 1005
 Ser Leu Lys Thr Met Gln Asp Lys Val Tyr Arg Gln Met Glu Leu Ala
 1010 1015 1020
 Arg Met Ile Val Ala Val Asp Glu His Asp Val Ala Glu Arg Val Ile
 1025 1030 1035 1040
 Asn Val His Phe Leu Pro Asp Ile Ile Gly Asn Leu Arg Ala Phe Ser
 1045 1050 1055
 Arg Gln Glu Phe Arg Cys Thr Arg Cys Asn Thr Lys Tyr Arg Arg Ile
 1060 1065 1070
 Pro Leu Val Gly Lys Cys Leu Lys Cys Gly Asn Lys Leu Thr Leu Thr
 1075 1080 1085

Val His Ser Ser Ser Ile Met Lys Tyr Leu Glu Leu Ser Lys Phe Leu
 1090 1095 1100

Cys Glu Asn Phe Asn Val Ser Ser Tyr Thr Lys Gln Arg Leu Met Leu
 1105 1110 1115 1120

Leu Glu Gln Glu Ile Lys Ser Met Phe Glu Asn Gly Thr Glu Lys Gln
 1125 1130 1135

Val Ser Ile Ser Asp Phe Val
 1140

(2) ANGABEN ZU SEQ ID NO: 28:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 1139 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: Methanococcus jannaschii

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 28:

Met Ile Val Met Val His Val Ala Cys Ser Glu Asn Met Lys Lys Tyr
 1 5 10 15

Phe Glu Asn Ile Val Asp Glu Val Lys Lys Ile Tyr Arg Ile Ala Glu
 20 25 30

Glu Cys Arg Lys Lys Gly Phe Asp Pro Thr Asp Glu Val Glu Ile Pro
 35 40 45

Leu Ala Ala Asp Met Ala Asp Arg Val Glu Gly Leu Val Gly Pro Lys
 50 55 60

Gly Val Ala Glu Arg Ile Arg Glu Leu Val Lys Glu Leu Gly Lys Glu
 65 70 75 80

Pro Ala Ala Leu Glu Ile Ala Lys Glu Ile Val Glu Gly Lys Phe Gly
 85 90 95

Asn Phe Asp Lys Glu Lys Lys Ala Glu Gln Ala Val Arg Thr Ala Leu
 100 105 110

Ala Val Leu Thr Glu Gly Ile Val Ala Ala Pro Leu Glu Gly Ile Ala
 115 120 125

Asp Val Lys Ile Lys Lys Asn Pro Asp Gly Thr Glu Tyr Leu Ala Ile
 130 135 140

Tyr Tyr Ala Gly Pro Ile Arg Ser Ala Gly Gly Thr Ala Gln Ala Leu
 145 150 155 160

Ser Val Leu Val Gly Asp Phe Val Arg Lys Ala Met Gly Leu Asp Arg
 165 170 175

Tyr Lys Pro Thr Glu Asp Glu Ile Glu Arg Tyr Val Glu Glu Val Glu
 180 185 190

Leu Tyr Gln Ser Glu Val Gly Ser Phe Gln Tyr Asn Pro Thr Ala Asp
 195 200 205

Glu Ile Arg Thr Ala Ile Arg Asn Ile Pro Ile Glu Ile Thr Gly Glu
 210 215 220

Ala Thr Asp Asp Val Glu Val Ser Gly His Arg Asp Leu Pro Arg Val
 225 230 235 240

Glu Thr Asn Gln Leu Arg Gly Gly Ala Leu Leu Val Leu Val Glu Gly
 245 250 255

Val Leu Leu Lys Ala Pro Lys Ile Leu Arg His Val Asp Lys Leu Gly
 260 265 270

Ile Glu Gly Trp Asp Trp Leu Lys Asp Leu Met Ser Lys Lys Glu Glu
 275 280 285

Lys Glu Glu Glu Lys Asp Glu Lys Val Asp Asp Glu Glu Ile Asp Glu
 290 295 300

Glu Glu Glu Glu Ile Ser Gly Tyr Trp Arg Asp Val Lys Ile Glu Ala
 305 310 315 320

Asn Lys Lys Phe Ile Ser Glu Val Ile Ala Gly Arg Pro Val Phe Ala
 325 330 335

His Pro Ser Lys Val Gly Gly Phe Arg Leu Arg Tyr Gly Arg Ser Arg
 340 345 350

Asn Thr Gly Phe Ala Thr Gln Gly Phe His Pro Ala Leu Met Tyr Leu
 355 360 365

Val Asp Glu Phe Met Ala Val Gly Thr Gln Leu Lys Thr Glu Arg Pro

370 375 380
 Gly Lys Ala Thr Cys Val Val Pro Val Asp Ser Ile Glu Pro Pro Ile
 385 390 395 400
 Val Lys Leu Lys Asn Gly Asp Val Ile Arg Val Asp Thr Ile Glu Lys
 405 410 415
 Ala Met Asp Val Arg Asn Arg Val Glu Glu Ile Leu Phe Leu Gly Asp
 420 425 430
 Val Leu Val Asn Tyr Gly Asp Phe Leu Glu Asn Asn His Pro Leu Leu
 435 440 445
 Pro Ser Cys Trp Cys Glu Glu Trp Tyr Glu Lys Ile Leu Ile Ala Asn
 450 455 460
 Asn Ile Glu Tyr Asp Lys Asp Phe Ile Lys Asn Pro Lys Pro Glu Glu
 465 470 475 480
 Ala Val Lys Phe Ala Leu Glu Thr Lys Thr Pro Leu His Pro Arg Phe
 485 490 495
 Thr Tyr His Trp His Asp Val Ser Lys Glu Asp Ile Ile Leu Leu Arg
 500 505 510
 Asn Trp Leu Leu Lys Gly Lys Glu Asp Ser Leu Glu Gly Lys Lys Val
 515 520 525
 Trp Ile Val Asp Leu Glu Ile Glu Glu Asp Lys Lys Ala Lys Arg Ile
 530 535 540
 Leu Glu Leu Ile Gly Cys Cys His Leu Val Arg Asn Lys Lys Val Ile
 545 550 555 560
 Ile Glu Glu Tyr Tyr Pro Leu Leu Tyr Ser Leu Gly Phe Asp Val Glu
 565 570 575
 Asn Lys Lys Asp Leu Val Glu Asn Ile Glu Lys Ile Leu Glu Ser Ala
 580 585 590
 Lys Asn Ser Met His Leu Ile Asn Leu Leu Ala Pro Phe Glu Val Arg
 595 600 605
 Arg Asn Thr Tyr Val Tyr Val Gly Ala Arg Met Gly Arg Pro Glu Lys
 610 615 620
 Ala Ala Pro Arg Lys Met Lys Pro Pro Val Asn Gly Leu Phe Pro Ile
 625 630 635 640

Gly Asn Ala Gly Gly Gln Val Arg Leu Ile Asn Lys Ala Val Glu Glu
 645 650 655

Asn Asn Thr Asp Asp Val Asp Val Ser Tyr Thr Arg Cys Pro Asn Cys
 660 665 670

Gly Lys Ile Ser Leu Tyr Arg Val Cys Pro Phe Cys Gly Thr Lys Val
 675 680 685

Glu Leu Asp Asn Phe Gly Arg Ile Lys Ala Pro Leu Lys Asp Tyr Trp
 690 695 700

Tyr Ala Ala Leu Lys Arg Leu Gly Ile Asn Lys Pro Gly Asp Val Lys
 705 710 715 720

Cys Ile Lys Gly Met Thr Ser Lys Gln Lys Ile Val Glu Pro Leu Glu
 725 730 735

Lys Ala Ile Leu Arg Ala Ile Asn Glu Val Tyr Val Phe Lys Asp Gly
 740 745 750

Thr Thr Arg Phe Asp Cys Thr Asp Val Pro Val Thr His Phe Lys Pro
 755 760 765

Asn Glu Ile Asn Val Thr Val Glu Lys Leu Arg Glu Leu Gly Tyr Asp
 770 775 780

Lys Asp Ile Tyr Gly Asn Glu Leu Val Asp Gly Glu Gln Val Val Glu
 785 790 795 800

Leu Lys Pro Gln Asp Val Ile Ile Pro Glu Ser Cys Ala Glu Tyr Phe
 805 810 815

Val Lys Val Ala Asn Phe Ile Asp Asp Leu Leu Glu Lys Phe Tyr Lys
 820 825 830

Val Glu Arg Phe Tyr Asn Val Lys Lys Lys Glu Asp Leu Ile Gly His
 835 840 845

Leu Val Ile Gly Met Ala Pro His Thr Ser Ala Gly Met Val Gly Arg
 850 855 860

Ile Ile Gly Tyr Thr Lys Ala Asn Val Gly Tyr Ala His Pro Tyr Phe
 865 870 875 880

His Ala Ala Lys Arg Arg Asn Cys Asp Gly Asp Glu Asp Ser Phe Phe
 885 890 895

Leu Leu Leu Asp Ala Phe Leu Asn Phe Ser Lys Lys Phe Leu Pro Asp

900 905 910
 Lys Arg Gly Gly Gln Met Asp Ala Pro Leu Val Leu Thr Thr Ile Leu
 915 920 925
 Asp Pro Lys Glu Val Asp Gly Glu Val His Asn Met Asp Thr Met Trp
 930 935 940
 Ser Tyr Pro Leu Glu Phe Tyr Glu Lys Thr Leu Glu Met Pro Ser Pro
 945 950 955 960
 Lys Glu Val Lys Glu Phe Met Glu Thr Val Glu Asp Arg Leu Gly Lys
 965 970 975
 Pro Glu Gln Tyr Glu Gly Ile Gly Tyr Thr His Glu Thr Ser Arg Ile
 980 985 990
 Asp Leu Gly Pro Lys Val Cys Ala Tyr Lys Thr Leu Gly Ser Met Leu
 995 1000 1005
 Glu Lys Thr Thr Ser Gln Leu Ser Val Ala Lys Lys Ile Arg Ala Thr
 1010 1015 1020
 Asp Glu Arg Asp Val Ala Glu Lys Val Ile Gln Ser His Phe Ile Pro
 1025 1030 1035 1040
 Asp Leu Ile Gly Asn Leu Arg Ala Phe Ser Arg Gln Ala Val Arg Cys
 1045 1050 1055
 Lys Cys Gly Ala Lys Tyr Arg Arg Ile Pro Leu Lys Gly Lys Cys Pro
 1060 1065 1070
 Lys Cys Gly Ser Asn Leu Ile Leu Thr Val Ser Lys Gly Ala Val Glu
 1075 1080 1085
 Lys Tyr Met Asp Val Ala Glu Lys Met Ala Glu Glu Tyr Asn Val Asn
 1090 1095 1100
 Asp Tyr Ile Lys Gln Arg Leu Lys Ile Ile Lys Glu Gly Ile Asn Ser
 1105 1110 1115 1120
 Ile Phe Glu Asn Glu Lys Ser Arg Gln Val Lys Leu Ser Asp Phe Phe
 1125 1130 1135
 Lys Ile Gly

(2) ANGABEN ZU SEQ ID NO: 29:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 1434 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: *Pyrococcus horikoshii*

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 29:

Met Val Leu Met Glu Leu Pro Lys Glu Met Glu Glu Tyr Phe Ser Met
1 5 10 15

Leu Gln Arg Glu Ile Asp Lys Ala Tyr Glu Ile Ala Lys Lys Ala Arg
 20 25 30

Ala Gln Gly Lys Asp Pro Ser Leu Asp Val Glu Ile Pro Gln Ala Ser
 35 40 45

Asp Met Ala Gly Arg Val Glu Ser Leu Val Gly Pro Pro Gly Val Ala
 50 55 60

Glu Arg Ile Arg Glu Leu Val Lys Glu Tyr Gly Lys Glu Ile Ala Ala
65 70 75 80

Leu Lys Ile Val Asp Glu Ile Ile Asp Gly Lys Phe Gly Asp Leu Gly
 85 90 95

Ser Lys Glu Lys Tyr Ala Glu Gln Ala Val Arg Thr Ala Leu Ala Ile
 100 105 110

Leu Thr Glu Gly Val Val Ser Ala Pro Ile Glu Gly Ile Ala Ser Val
 115 120 125

Lys Ile Lys Arg Asn Thr Trp Ser Asp Asn Ser Glu Tyr Leu Ala Leu
 130 135 140

Tyr Tyr Ala Gly Pro Ile Arg Ser Ser Gly Gly Thr Ala Gln Ala Leu
145 150 155 160

Ser Val Leu Val Gly Asp Tyr Val Arg Arg Lys Leu Gly Leu Asp Arg
 165 170 175

Phe Lys Pro Ser Glu Lys His Ile Glu Arg Met Val Glu Glu Val Asp

180	185	190
Leu Tyr His Arg Thr Val Ser Arg Leu Gln Tyr His Pro Ser Pro Glu		
195	200	205
Glu Val Arg Leu Ala Met Arg Asn Ile Pro Ile Glu Ile Thr Gly Glu		
210	215	220
Ala Thr Asp Glu Val Glu Val Ser His Arg Asp Ile Pro Gly Val Glu		
225	230	235
Thr Asn Gln Leu Arg Gly Gly Ala Ile Leu Val Leu Ala Glu Gly Val		
245	250	255
Leu Gln Lys Ala Lys Lys Leu Val Lys Tyr Ile Asp Lys Met Gly Ile		
260	265	270
Glu Gly Trp Glu Trp Leu Lys Glu Phe Val Glu Ala Lys Glu Lys Gly		
275	280	285
Glu Glu Ile Glu Glu Glu Gly Ser Ala Glu Ser Thr Val Glu Glu Thr		
290	295	300
Lys Val Glu Val Asp Met Gly Phe Tyr Tyr Ser Leu Tyr Gln Lys Phe		
305	310	315
Lys Ser Glu Ile Ala Pro Asn Asp Lys Tyr Ala Lys Glu Ile Ile Gly		
325	330	335
Gly Arg Pro Leu Phe Ser Asp Pro Ser Arg Asn Gly Gly Phe Arg Leu		
340	345	350
Arg Tyr Gly Arg Ser Arg Val Ser Gly Phe Ala Thr Trp Gly Ile Asn		
355	360	365
Pro Ala Thr Met Ile Leu Val Asp Glu Phe Leu Ala Ile Gly Thr Gln		
370	375	380
Leu Lys Thr Glu Arg Pro Gly Lys Gly Ala Val Val Thr Pro Val Thr		
385	390	395
Thr Ile Glu Gly Pro Ile Val Lys Leu Lys Asp Gly Ser Val Val Lys		
405	410	415
Val Asp Asp Tyr Lys Leu Ala Leu Lys Ile Arg Asp Glu Val Glu Glu		
420	425	430
Ile Leu Tyr Leu Gly Asp Ala Val Ile Ala Phe Gly Asp Phe Val Glu		
435	440	445

Asn Asn Gln Thr Leu Leu Pro Ala Asn Tyr Cys Glu Glu Trp Trp Ile
 450 455 460

Leu Glu Phe Thr Lys Ala Leu Asn Glu Ile Tyr Glu Val Glu Leu Lys
 465 470 475 480

Pro Phe Glu Val Asn Ser Ser Glu Asp Leu Glu Glu Ala Ala Asp Tyr
 485 490 495

Leu Glu Val Asp Ile Glu Phe Leu Lys Glu Leu Leu Lys Asp Pro Leu
 500 505 510

Arg Thr Lys Pro Pro Val Glu Leu Ala Ile His Phe Ser Glu Ile Leu
 515 520 525

Gly Ile Pro Leu His Pro Tyr Tyr Thr Leu Tyr Trp Asn Ser Val Lys
 530 535 540

Pro Glu Gln Val Glu Lys Leu Trp Arg Val Leu Lys Glu His Ala His
 545 550 555 560

Ile Asp Trp Asp Asn Phe Arg Gly Ile Lys Phe Ala Arg Arg Ile Val
 565 570 575

Ile Pro Leu Glu Lys Leu Arg Asp Ser Lys Arg Ala Leu Glu Leu Leu
 580 585 590

Gly Leu Pro His Lys Val Glu Gly Lys Asn Val Ile Val Asp Tyr Pro
 595 600 605

Trp Ala Ala Ala Leu Leu Thr Pro Leu Gly Asn Leu Glu Trp Glu Phe
 610 615 620

Arg Ala Lys Pro Leu His Thr Thr Ile Asp Ile Ile Asn Glu Asn Asn
 625 630 635 640

Glu Ile Lys Leu Arg Asp Arg Gly Ile Ser Trp Ile Gly Ala Arg Met
 645 650 655

Gly Arg Pro Glu Lys Ala Lys Glu Arg Lys Met Lys Pro Pro Val Gln
 660 665 670

Val Leu Phe Pro Ile Gly Leu Ala Gly Gly Ser Ser Arg Asp Ile Lys
 675 680 685

Lys Ala Ala Glu Glu Gly Lys Val Ala Glu Val Glu Ile Ala Leu Phe
 690 695 700

Lys Cys Pro Lys Cys Gly His Val Gly Pro Glu His Ile Cys Pro Asn

705 710 715 720
 Cys Gly Thr Arg Lys Glu Leu Ile Trp Val Cys Pro Arg Cys Asn Ala
 725 730 735
 Glu Tyr Pro Glu Ser Gln Ala Ser Gly Tyr Asn Tyr Thr Cys Pro Lys
 740 745 750
 Cys Asn Val Lys Leu Lys Pro Tyr Ala Lys Arg Lys Ile Lys Pro Ser
 755 760 765
 Glu Leu Leu Lys Arg Ala Met Asp Asn Val Lys Val Tyr Gly Ile Asp
 770 775 780
 Lys Leu Lys Gly Val Met Gly Met Thr Ser Gly Trp Lys Met Pro Glu
 785 790 795 800
 Pro Leu Glu Lys Gly Leu Leu Arg Ala Lys Asn Asp Val Tyr Val Phe
 805 810 815
 Lys Asp Gly Thr Ile Arg Phe Asp Ala Thr Asp Ala Pro Ile Thr His
 820 825 830
 Phe Arg Pro Arg Glu Ile Gly Val Ser Val Glu Lys Leu Arg Glu Leu
 835 840 845
 Gly Tyr Thr His Asp Phe Glu Gly Asn Pro Leu Val Ser Glu Asp Gln
 850 855 860
 Ile Val Glu Leu Lys Pro Gln Asp Ile Ile Leu Ser Lys Glu Ala Gly
 865 870 875 880
 Lys Tyr Leu Leu Lys Val Ala Lys Phe Val Asp Asp Leu Leu Glu Lys
 885 890 895
 Phe Tyr Gly Leu Pro Arg Phe Tyr Asn Ala Glu Lys Met Glu Asp Leu
 900 905 910
 Ile Gly His Leu Val Ile Gly Leu Ala Pro His Thr Ser Ala Gly Ile
 915 920 925
 Val Gly Arg Ile Ile Gly Phe Val Asp Ala Leu Val Gly Tyr Ala His
 930 935 940
 Pro Tyr Phe His Ala Ala Lys Arg Arg Asn Cys Phe Pro Gly Asp Thr
 945 950 955 960
 Arg Ile Leu Val Gln Ile Asn Gly Thr Pro Gln Arg Val Thr Leu Lys
 965 970 975

Glu Leu Tyr Glu Leu Phe Asp Glu Glu His Tyr Glu Ser Met Val Tyr
 980 985 990

Val Arg Lys Lys Pro Lys Val Asp Ile Lys Val Tyr Ser Phe Asn Pro
 995 1000 1005

Glu Glu Gly Lys Val Val Leu Thr Asp Ile Glu Glu Val Ile Lys Ala
 1010 1015 1020

Pro Ala Thr Asp His Leu Ile Arg Phe Glu Leu Glu Leu Gly Ser Ser
 1025 1030 1035 1040

Phe Glu Thr Thr Val Asp His Pro Val Leu Val Tyr Glu Asn Gly Lys
 1045 1050 1055

Phe Val Glu Lys Arg Ala Phe Glu Val Arg Glu Gly Asn Ile Ile Ile
 1060 1065 1070

Ile Ile Asp Glu Ser Thr Leu Glu Pro Leu Lys Val Ala Val Lys Lys
 1075 1080 1085

Ile Glu Phe Ile Glu Pro Pro Glu Asp Phe Val Phe Ser Leu Asn Ala
 1090 1095 1100

Lys Lys Tyr His Thr Val Ile Ile Asn Glu Asn Ile Val Thr His Gln
 1105 1110 1115 1120

Cys Asp Gly Asp Glu Asp Ala Val Met Leu Leu Leu Asp Ala Leu Leu
 1125 1130 1135

Asn Phe Ser Arg Tyr Tyr Leu Pro Glu Lys Arg Gly Gly Lys Met Asp
 1140 1145 1150

Ala Pro Leu Val Ile Thr Thr Arg Leu Asp Pro Arg Glu Val Asp Ser
 1155 1160 1165

Glu Val His Asn Met Asp Ile Val Arg Tyr Tyr Pro Leu Glu Phe Tyr
 1170 1175 1180

Glu Ala Thr Tyr Glu Leu Lys Ser Pro Lys Glu Leu Val Gly Val Ile
 1185 1190 1195 1200

Glu Arg Val Glu Asp Arg Leu Gly Lys Pro Glu Met Tyr Tyr Gly Leu
 1205 1210 1215

Lys Phe Thr His Asp Thr Asp Asp Ile Ala Leu Gly Pro Lys Met Ser
 1220 1225 1230

Leu Tyr Lys Gln Leu Gly Asp Met Glu Glu Lys Val Lys Arg Gln Leu

1235	1240	1245
Asp Val Ala Arg Arg Ile Arg Ala Val Asp Glu His Lys Val Ala Glu		
1250	1255	1260
Thr Ile Leu Asn Ser His Leu Ile Pro Asp Leu Arg Gly Asn Leu Arg		
1265	1270	1275
		1280
Ser Phe Thr Arg Gln Glu Phe Arg Cys Val Lys Cys Asn Thr Lys Phe		
1285	1290	1295
Arg Arg Pro Pro Leu Asp Gly Lys Cys Pro Ile Cys Gly Gly Lys Ile		
1300	1305	1310
Val Leu Thr Val Ser Lys Gly Ala Ile Glu Lys Tyr Leu Gly Thr Ala		
1315	1320	1325
Lys Met Leu Val Thr Glu Tyr Lys Val Lys Asn Tyr Thr Arg Gln Arg		
1330	1335	1340
Ile Cys Leu Thr Glu Arg Asp Ile Asp Ser Leu Phe Glu Thr Val Phe		
1345	1350	1355
		1360
Pro Glu Thr Gln Leu Thr Leu Leu Val Asn Pro Asn Asp Ile Cys Gln		
1365	1370	1375
Arg Ile Ile Met Glu Arg Thr Gly Gly Ser Lys Lys Ser Gly Leu Leu		
1380	1385	1390
Glu Asn Phe Ala Asn Gly Tyr Asn Lys Gly Lys Lys Glu Glu Met Pro		
1395	1400	1405
Lys Lys Gln Arg Lys Lys Glu Gln Glu Lys Ser Lys Lys Arg Lys Val		
1410	1415	1420
Ile Ser Leu Asp Asp Phe Phe Ser Arg Lys		
1425	1430	

(2) ANGABEN ZU SEQ ID NO: 30:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 1092 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

(A) ORGANISMUS: *Methanobacterium thermoautotrophicum*

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 30:

```

Met Met Asp Tyr Phe Asn Glu Leu Glu Arg Glu Thr Glu Arg Leu Tyr
 1           5           10          15

Glu Ile Ala Arg Lys Ala Arg Ala Arg Gly Leu Asp Val Ser Thr Thr
      20           25           30

Pro Glu Ile Pro Leu Ala Lys Asp Leu Ala Glu Arg Val Glu Gly Leu
      35           40           45

Val Gly Pro Glu Gly Ile Ala Arg Arg Ile Lys Glu Leu Glu Gly Asp
      50           55           60

Arg Gly Arg Glu Glu Val Ala Phe Gln Ile Ala Ala Glu Ile Ala Ser
      65           70           75           80

Gln Ala Val Pro Asp Asp Asp Pro Glu Glu Arg Glu Lys Leu Ala Asp
      85           90           95

Gln Ala Leu Arg Thr Ala Leu Ala Ile Leu Thr Glu Gly Val Val Ala
      100          105          110

Ala Pro Leu Glu Gly Ile Ala Arg Val Arg Ile Lys Glu Asn Phe Asp
      115          120          125

Lys Ser Arg Tyr Leu Ala Val Tyr Phe Ala Gly Pro Ile Arg Ser Ala
      130          135          140

Gly Gly Thr Ala Ala Ala Leu Ser Val Leu Ile Ala Asp Tyr Ile Arg
      145          150          155          160

Leu Ala Val Gly Leu Asp Arg Tyr Lys Pro Val Glu Arg Glu Ile Glu
      165          170          175

Arg Tyr Val Glu Glu Val Glu Leu Tyr Glu Ser Glu Val Thr Asn Leu
      180          185          190

Gln Tyr Ser Pro Lys Pro Asp Glu Val Arg Leu Ala Ala Ser Lys Ile
      195          200          205

Pro Val Glu Val Thr Gly Glu Pro Thr Asp Lys Val Glu Val Ser His
      210          215          220

Arg Asp Leu Glu Arg Val Glu Thr Asn Asn Ile Arg Gly Gly Ala Leu

```

225 230 235 240
Leu Ala Met Val Glu Gly Val Ile Gln Lys Ala Pro Lys Val Leu Lys
 245 250 255
Tyr Ala Lys Gln Leu Lys Leu Glu Gly Trp Asp Trp Leu Glu Lys Phe
 260 265 270
Ser Lys Ala Pro Lys Lys Gly Glu Gly Glu Glu Lys Val Val Val Lys
 275 280 285
Ala Asp Ser Lys Tyr Val Glu Asp Ile Ile Gly Gly Arg Pro Val Leu
 290 295 300
Ala Tyr Pro Ser Glu Lys Gly Ala Phe Arg Leu Arg Tyr Gly Arg Ala
305 310 315 320
Arg Asn Thr Gly Leu Ala Ala Met Gly Val His Pro Ala Thr Met Glu
 325 330 335
Leu Leu Gln Phe Leu Ala Val Gly Thr Gln Met Lys Ile Glu Arg Pro
 340 345 350
Gly Lys Gly Asn Cys Val Val Pro Val Asp Thr Ile Asp Gly Pro Val
 355 360 365
Val Lys Leu Arg Asn Gly Asp Val Ile Arg Ile Glu Asp Ala Glu Thr
 370 375 380
Ala Ser Arg Val Arg Ser Glu Val Glu Glu Ile Leu Phe Leu Gly Asp
385 390 395 400
Met Leu Val Ala Phe Gly Glu Phe Leu Arg Asn Asn His Val Leu Met
 405 410 415
Pro Ala Gly Trp Cys Glu Glu Trp Trp Ile Gln Thr Ile Leu Ser Ser
 420 425 430
Pro Lys Tyr Pro Gly Asp Asp Pro Leu Asn Leu Ser Tyr Tyr Arg Thr
 435 440 445
Arg Trp Asn Glu Leu Glu Val Ser Ala Gly Asp Ala Phe Arg Ile Ser
 450 455 460
Glu Glu Tyr Asp Val Pro Leu His Pro Arg Tyr Thr Tyr Phe Tyr His
465 470 475 480
Asp Val Thr Val Arg Glu Leu Asn Met Leu Arg Glu Trp Leu Asn Thr
 485 490 495

Ser Gln Leu Glu Asp Glu Leu Val Leu Glu Leu Arg Pro Glu Lys Arg
 500 505 510

Ile Leu Glu Ile Leu Gly Val Pro His Arg Val Lys Asp Ser Arg Val
 515 520 525

Val Ile Gly His Asp Asp Ala His Ala Leu Ile Lys Thr Leu Arg Lys
 530 535 540

Pro Leu Glu Asp Ser Ser Asp Thr Val Glu Ala Leu Asn Arg Val Ser
 545 550 555 560

Pro Val Arg Ile Met Lys Lys Ala Pro Thr Tyr Ile Gly Thr Arg Val
 565 570 575

Gly Arg Pro Glu Lys Thr Lys Glu Arg Lys Met Arg Pro Ala Pro His
 580 585 590

Val Leu Phe Pro Ile Gly Lys Tyr Gly Gly Ser Arg Arg Asn Ile Pro
 595 600 605

Asp Ala Ala Lys Lys Gly Ser Ile Thr Val Glu Ile Gly Arg Ala Thr
 610 615 620

Cys Pro Ser Cys Arg Val Ser Ser Met Gln Ser Ile Cys Pro Ser Cys
 625 630 635 640

Gly Ser Arg Thr Val Ile Gly Glu Pro Gly Lys Arg Asn Ile Asn Leu
 645 650 655

Ala Ala Leu Leu Lys Arg Ala Ala Glu Asn Val Ser Val Arg Lys Leu
 660 665 670

Asp Glu Ile Lys Gly Val Glu Gly Met Ile Ser Ala Glu Lys Phe Pro
 675 680 685

Glu Pro Leu Glu Lys Gly Ile Leu Arg Ala Lys Asn Asp Val Tyr Thr
 690 695 700

Phe Lys Asp Ala Thr Ile Arg His Asp Ser Thr Asp Leu Pro Leu Thr
 705 710 715 720

His Phe Thr Pro Arg Glu Val Gly Val Ser Val Glu Arg Leu Arg Glu
 725 730 735

Leu Gly Tyr Thr Arg Asp Cys Tyr Gly Asp Glu Leu Glu Asp Glu Asp
 740 745 750

Gln Ile Leu Glu Leu Arg Val Gln Asp Val Val Ile Ser Glu Asp Cys

755	760	765
Ala Asp Tyr Leu Val Arg Val Ala Asn Phe Val Asp Asp Leu Leu Glu		
770	775	780
Arg Phe Tyr Asp Leu Glu Arg Phe Tyr Asn Val Lys Thr Arg Glu Asp		
785	790	795 800
Leu Val Gly His Leu Ile Ala Gly Leu Ala Pro His Thr Ser Ala Ala		
805	810	815
Val Leu Gly Arg Ile Ile Gly Phe Thr Gly Ala Ser Ala Cys Tyr Ala		
820	825	830
His Pro Tyr Phe His Ser Ala Lys Arg Arg Asn Cys Asp Ser Asp Glu		
835	840	845
Asp Ser Val Met Leu Leu Leu Asp Ala Leu Leu Asn Phe Ser Lys Ser		
850	855	860
Tyr Leu Pro Ser Ser Arg Gly Gly Ser Met Asp Ala Pro Leu Val Leu		
865	870	875 880
Ser Thr Arg Ile Asp Pro Glu Glu Ile Asp Asp Glu Ser His Asn Ile		
885	890	895
Asp Thr Met Asp Met Ile Pro Leu Glu Val Tyr Glu Arg Ser Phe Asp		
900	905	910
His Pro Arg Pro Ser Glu Val Leu Asp Val Ile Asp Asn Val Glu Lys		
915	920	925
Arg Leu Gly Lys Pro Glu Gln Tyr Thr Gly Leu Met Phe Ser His Asn		
930	935	940
Thr Ser Arg Ile Asp Glu Gly Pro Lys Val Cys Leu Tyr Lys Leu Leu		
945	950	955 960
Pro Thr Met Lys Glu Lys Val Glu Ser Gln Ile Thr Leu Ala Glu Lys		
965	970	975
Ile Arg Ala Val Asp Gln Arg Ser Val Val Glu Gly Val Leu Met Ser		
980	985	990
His Phe Leu Pro Asp Met Met Gly Asn Ile Arg Ala Phe Ser Arg Gln		
995	1000	1005
Lys Val Arg Cys Thr Lys Cys Asn Arg Lys Tyr Arg Arg Ile Pro Leu		
1010	1015	1020

Ser Gly Glu Cys Arg Cys Gly Gly Asn Leu Val Leu Thr Val Ser Lys
 1025 1030 1035 1040

Gly Ser Val Ile Lys Tyr Leu Glu Ile Ser Lys Glu Leu Ala Ser Arg
 1045 1050 1055

Tyr Pro Ile Asp Pro Tyr Leu Met Gln Arg Ile Glu Ile Leu Glu Tyr
 1060 1065 1070

Gly Val Asn Ser Leu Phe Glu Ser Asp Arg Ser Lys Gln Ser Ser Leu
 1075 1080 1085

Asp Val Phe Leu
 1090

(2) ANGABEN ZU SEQ ID NO: 31:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 1263 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: Pyrococcus furiosus

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 31:

Met Glu Leu Pro Lys Glu Ile Glu Glu Tyr Phe Glu Met Leu Gln Arg
 1 5 10 15

Glu Ile Asp Lys Ala Tyr Glu Ile Ala Lys Lys Ala Arg Ser Gln Gly
 20 25 30

Lys Asp Pro Ser Thr Asp Val Glu Ile Pro Gln Ala Thr Asp Met Ala
 35 40 45

Gly Arg Val Glu Ser Leu Val Gly Pro Pro Gly Val Ala Gln Arg Ile
 50 55 60

Arg Glu Leu Leu Lys Glu Tyr Asp Lys Glu Ile Val Ala Leu Lys Ile
 65 70 75 80

Val Asp Glu Ile Ile Glu Gly Lys Phe Gly Asp Phe Gly Ser Lys Glu
 85 90 95

Lys Tyr Ala Glu Gln Ala Val Arg Thr Ala Leu Ala Ile Leu Thr Glu
100 105 110

Gly Ile Val Ser Ala Pro Leu Glu Gly Ile Ala Asp Val Lys Ile Lys
115 120 125

Arg Asn Thr Trp Ala Asp Asn Ser Glu Tyr Leu Ala Leu Tyr Tyr Ala
130 135 140

Gly Pro Ile Arg Ser Ser Gly Gly Thr Ala Gln Ala Leu Ser Val Leu
145 150 155 160

Val Gly Asp Tyr Val Arg Arg Lys Leu Gly Leu Asp Arg Phe Lys Pro
165 170 175

Ser Gly Lys His Ile Glu Arg Met Val Glu Glu Val Asp Leu Tyr His
180 185 190

Arg Ala Val Ser Arg Leu Gln Tyr His Pro Ser Pro Asp Glu Val Arg
195 200 205

Leu Ala Met Arg Asn Ile Pro Ile Glu Ile Thr Gly Glu Ala Thr Asp
210 215 220

Asp Val Glu Val Ser His Arg Asp Val Glu Gly Val Glu Thr Asn Gln
225 230 235 240

Leu Arg Gly Gly Ala Ile Leu Val Leu Ala Glu Gly Val Leu Gln Lys
245 250 255

Ala Lys Lys Leu Val Lys Tyr Ile Asp Lys Met Gly Ile Asp Gly Trp
260 265 270

Glu Trp Leu Lys Glu Phe Val Glu Ala Lys Glu Lys Gly Glu Glu Ile
275 280 285

Glu Glu Ser Glu Ser Lys Ala Glu Glu Ser Lys Val Glu Thr Arg Val
290 295 300

Glu Val Glu Lys Gly Phe Tyr Tyr Lys Leu Tyr Glu Lys Phe Arg Ala
305 310 315 320

Glu Ile Ala Pro Ser Glu Lys Tyr Ala Lys Glu Ile Ile Gly Gly Arg
325 330 335

Pro Leu Phe Ala Gly Pro Ser Glu Asn Gly Gly Phe Arg Leu Arg Tyr
340 345 350

Gly Arg Ser Arg Val Ser Gly Phe Ala Thr Trp Ser Ile Asn Pro Ala

355	360	365
Thr Met Val Leu Val Asp Glu Phe Leu Ala Ile Gly Thr Gln Met Lys		
370	375	380
Thr Glu Arg Pro Gly Lys Gly Ala Val Val Thr Pro Ala Thr Thr Ala		
385	390	395
		400
Glu Gly Pro Ile Val Lys Leu Lys Asp Gly Ser Val Val Arg Val Asp		
405	410	415
Asp Tyr Asn Leu Ala Leu Lys Ile Arg Asp Glu Val Glu Glu Ile Leu		
420	425	430
Tyr Leu Gly Asp Ala Ile Ile Ala Phe Gly Asp Phe Val Glu Asn Asn		
435	440	445
Gln Thr Leu Leu Pro Ala Asn Tyr Val Glu Glu Trp Trp Ile Gln Glu		
450	455	460
Phe Val Lys Ala Val Asn Glu Ala Tyr Glu Val Glu Leu Arg Pro Phe		
465	470	475
		480
Glu Glu Asn Pro Arg Glu Ser Val Glu Glu Ala Ala Glu Tyr Leu Glu		
485	490	495
Val Asp Pro Glu Phe Leu Ala Lys Met Leu Tyr Asp Pro Leu Arg Val		
500	505	510
Lys Pro Pro Val Glu Leu Ala Ile His Phe Ser Glu Ile Leu Glu Ile		
515	520	525
Pro Leu His Pro Tyr Tyr Thr Leu Tyr Trp Asn Thr Val Asn Pro Lys		
530	535	540
Asp Val Glu Arg Leu Trp Gly Val Leu Lys Asp Lys Ala Thr Ile Glu		
545	550	555
		560
Trp Gly Thr Phe Arg Gly Ile Lys Phe Ala Lys Lys Ile Glu Ile Ser		
565	570	575
Leu Asp Asp Leu Gly Ser Leu Lys Arg Thr Leu Glu Leu Leu Gly Leu		
580	585	590
Pro His Thr Val Arg Glu Gly Ile Val Val Val Asp Tyr Pro Trp Ser		
595	600	605
Ala Ala Leu Leu Thr Pro Leu Gly Asn Leu Glu Trp Glu Phe Lys Ala		
610	615	620

Lys Pro Phe Tyr Thr Val Ile Asp Ile Ile Asn Glu Asn Asn Gln Ile
625 630 635 640

Lys Leu Arg Asp Arg Gly Ile Ser Trp Ile Gly Ala Arg Met Gly Arg
 645 650 655

Pro Glu Lys Ala Lys Glu Arg Lys Met Lys Pro Pro Val Gln Val Leu
 660 665 670

Phe Pro Ile Gly Leu Ala Gly Gly Ser Ser Arg Asp Ile Lys Lys Ala
 675 680 685

Ala Glu Glu Gly Lys Ile Ala Glu Val Glu Ile Ala Phe Phe Lys Cys
 690 695 700

Pro Lys Cys Gly His Val Gly Pro Glu Thr Leu Cys Pro Glu Cys Gly
705 710 715 720

Ile Arg Lys Glu Leu Ile Trp Thr Cys Pro Lys Cys Gly Ala Glu Tyr
 725 730 735

Thr Asn Ser Gln Ala Glu Gly Tyr Ser Tyr Ser Cys Pro Lys Cys Asn
 740 745 750

Val Lys Leu Lys Pro Phe Thr Lys Arg Lys Ile Lys Pro Ser Glu Leu
 755 760 765

Leu Asn Arg Ala Met Glu Asn Val Lys Val Tyr Gly Val Asp Lys Leu
 770 775 780

Lys Gly Val Met Gly Met Thr Ser Gly Trp Lys Ile Ala Glu Pro Leu
785 790 795 800

Glu Lys Gly Leu Leu Arg Ala Lys Asn Glu Val Tyr Val Phe Lys Asp
 805 810 815

Gly Thr Ile Arg Phe Asp Ala Thr Asp Ala Pro Ile Thr His Phe Arg
 820 825 830

Pro Arg Glu Ile Gly Val Ser Val Glu Lys Leu Arg Glu Leu Gly Tyr
 835 840 845

Thr His Asp Phe Glu Gly Lys Pro Leu Val Ser Glu Asp Gln Ile Val
 850 855 860

Glu Leu Lys Pro Gln Asp Val Ile Leu Ser Lys Glu Ala Gly Lys Tyr
865 870 875 880

Leu Leu Arg Val Ala Arg Phe Val Asp Asp Leu Leu Glu Lys Phe Tyr

	885	890	895	
Gly Leu Pro Arg Phe Tyr Asn Ala Glu Lys Met Glu Asp Leu Ile Gly				
	900	905	910	
His Leu Val Ile Gly Leu Ala Pro His Thr Ser Ala Gly Ile Val Gly				
	915	920	925	
Arg Ile Ile Gly Phe Val Asp Ala Leu Val Gly Tyr Ala His Pro Tyr				
	930	935	940	
Phe His Ala Ala Lys Arg Arg Asn Cys Asp Gly Asp Glu Asp Ser Val				
	945	950	955	960
Met Leu Leu Leu Asp Ala Leu Leu Asn Phe Ser Arg Tyr Tyr Leu Pro				
	965	970	975	
Glu Lys Arg Gly Gly Lys Met Asp Ala Pro Leu Val Ile Thr Thr Arg				
	980	985	990	
Leu Asp Pro Arg Glu Val Asp Ser Glu Val His Asn Met Asp Val Val				
	995	1000	1005	
Arg Tyr Tyr Pro Leu Glu Phe Tyr Glu Ala Thr Tyr Glu Leu Lys Ser				
	1010	1015	1020	
Pro Lys Glu Leu Val Arg Val Ile Glu Gly Val Glu Asp Arg Leu Gly				
	1025	1030	1035	1040
Lys Pro Glu Met Tyr Tyr Gly Ile Lys Phe Thr His Asp Thr Asp Asp				
	1045	1050	1055	
Ile Ala Leu Gly Pro Lys Met Ser Leu Tyr Lys Gln Leu Gly Asp Met				
	1060	1065	1070	
Glu Glu Lys Val Lys Arg Gln Leu Thr Leu Ala Glu Arg Ile Arg Ala				
	1075	1080	1085	
Val Asp Gln His Tyr Val Ala Glu Thr Ile Leu Asn Ser His Leu Ile				
	1090	1095	1100	
Pro Asp Leu Arg Gly Asn Leu Arg Ser Phe Thr Arg Gln Glu Phe Arg				
	1105	1110	1115	1120
Cys Val Lys Cys Asn Thr Lys Tyr Arg Arg Pro Pro Leu Asp Gly Lys				
	1125	1130	1135	
Cys Pro Val Cys Gly Gly Lys Ile Val Leu Thr Val Ser Lys Gly Ala				
	1140	1145	1150	

Ile Glu Lys Tyr Leu Gly Thr Ala Lys Met Leu Val Ala Asn Tyr Asn
 1155 1160 1165

Val Lys Pro Tyr Thr Arg Gln Arg Ile Cys Leu Thr Glu Lys Asp Ile
 1170 1175 1180

Asp Ser Leu Phe Glu Tyr Leu Phe Pro Glu Ala Gln Leu Thr Leu Ile
 1185 1190 1195 1200

Val Asp Pro Asn Asp Ile Cys Met Lys Met Ile Lys Glu Arg Thr Gly
 1205 1210 1215

Glu Thr Val Gln Gly Gly Leu Leu Glu Asn Phe Asn Ser Ser Gly Asn
 1220 1225 1230

Asn Gly Lys Lys Ile Glu Lys Lys Glu Lys Lys Ala Lys Glu Lys Pro
 1235 1240 1245

Lys Lys Lys Lys Val Ile Ser Leu Asp Asp Phe Phe Ser Lys Arg
 1250 1255 1260

(2) ANGABEN ZU SEQ ID NO: 32:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 363 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: Homo sapiens

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 32:

Met Gln Ala Phe Leu Lys Gly Thr Ser Ile Ser Thr Lys Pro Pro Leu
 1 5 10 15

Thr Lys Asp Arg Gly Val Ala Ala Ser Ala Gly Ser Ser Gly Glu Asn
 20 25 30

Lys Lys Ala Lys Pro Val Pro Trp Val Glu Lys Tyr Arg Pro Lys Cys
 35 40 45

Val Asp Glu Val Ala Phe Gln Glu Glu Val Val Ala Val Leu Lys Lys
 50 55 60

Ser Leu Glu Gly Ala Asp Leu Pro Asn Leu Leu Phe Tyr Gly Pro Pro
65 70 75 80

Gly Thr Gly Lys Thr Ser Thr Ile Leu Ala Ala Ala Arg Glu Leu Phe
85 90 95

Gly Pro Glu Leu Phe Arg Leu Arg Val Leu Glu Leu Asn Ala Ser Asp
100 105 110

Glu Arg Gly Ile Gln Val Val Arg Glu Lys Val Lys Asn Phe Ala Gln
115 120 125

Leu Thr Val Ser Gly Ser Arg Ser Asp Gly Lys Pro Cys Pro Pro Phe
130 135 140

Lys Ile Val Ile Leu Asp Glu Ala Asp Ser Met Thr Ser Ala Ala Gln
145 150 155 160

Ala Ala Leu Arg Arg Thr Met Glu Lys Glu Ser Lys Thr Thr Arg Phe
165 170 175

Cys Leu Ile Cys Asn Tyr Val Ser Arg Ile Ile Glu Pro Leu Thr Ser
180 185 190

Arg Cys Ser Lys Phe Arg Phe Lys Pro Leu Ser Asp Lys Ile Gln Gln
195 200 205

Gln Arg Leu Leu Asp Ile Ala Lys Lys Glu Asn Val Lys Ile Ser Asp
210 215 220

Glu Gly Ile Ala Tyr Leu Val Lys Val Ser Glu Gly Asp Leu Arg Lys
225 230 235 240

Ala Ile Thr Phe Leu Gln Ser Ala Thr Arg Leu Thr Gly Gly Lys Glu
245 250 255

Ile Thr Glu Lys Val Ile Thr Asp Ile Ala Gly Val Ile Pro Ala Glu
260 265 270

Lys Ile Asp Gly Val Phe Ala Ala Cys Gln Ser Gly Ser Phe Asp Lys
275 280 285

Leu Glu Ala Val Val Lys Asp Leu Ile Asp Glu Gly His Ala Ala Thr
290 295 300

Gln Leu Val Asn Gln Leu His Asp Val Val Val Glu Asn Asn Leu Ser
305 310 315 320

Asp Lys Gln Lys Ser Ile Ile Thr Glu Lys Leu Ala Glu Val Asp Lys

325 330 335
 Cys Leu Ala Asp Gly Ala Asp Glu His Leu Gln Leu Ile Ser Leu Cys
 340 345 350
 Ala Thr Val Met Gln Gln Leu Ser Gln Asn Cys
 355 360

(2) ANGABEN ZU SEQ ID NO: 33:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 329 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: Homo sapiens

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 33:

Asn Leu Val Gln Cys Gly Asp Phe Pro His Leu Leu Val Tyr Gly Pro
 1 5 10 15
 Ser Gly Ala Gly Lys Lys Thr Arg Ile Met Cys Ile Leu Arg Glu Leu
 20 25 30
 Tyr Gly Val Gly Val Glu Lys Leu Arg Ile Glu His Gln Thr Ile Thr
 35 40 45
 Thr Pro Ser Lys Lys Lys Ile Glu Ile Ser Thr Ile Ala Ser Asn Tyr
 50 55 60
 His Leu Glu Val Asn Pro Ser Asp Ala Gly Asn Ser Asp Arg Val Val
 65 70 75 80
 Ile Gln Glu Met Leu Lys Thr Val Ala Gln Ser Gln Gln Leu Glu Thr
 85 90 95
 Asn Ser Gln Arg Asp Phe Lys Val Val Leu Leu Thr Glu Val Asp Lys
 100 105 110
 Leu Thr Lys Asp Ala Gln His Ala Leu Arg Arg Thr Met Glu Lys Tyr
 115 120 125

Met Ser Thr Cys Arg Leu Ile Leu Cys Cys Asn Ser Thr Ser Lys Val
 130 135 140

Ile Pro Pro Ile Arg Ser Arg Cys Leu Ala Val Arg Val Pro Ala Pro
 145 150 155 160

Ser Ile Glu Asp Ile Cys His Val Leu Ser Thr Val Cys Lys Lys Glu
 165 170 175

Gly Leu Asn Leu Pro Ser Gln Leu Ala His Arg Leu Ala Glu Lys Ser
 180 185 190

Cys Arg Asn Leu Arg Lys Ala Leu Leu Met Cys Glu Ala Cys Arg Val
 195 200 205

Gln Gln Tyr Pro Phe Thr Ala Asp Gln Glu Ile Pro Glu Thr Asp Trp
 210 215 220

Glu Val Tyr Leu Arg Glu Thr Ala Asn Ala Ile Val Ser Gln Gln Thr
 225 230 235 240

Pro Gln Arg Leu Leu Glu Val Arg Gly Arg Leu Tyr Glu Leu Leu Thr
 245 250 255

His Cys Ile Pro Pro Glu Ile Ile Met Lys Gly Leu Leu Ser Glu Leu
 260 265 270

Leu His Asn Cys Asp Gly Gln Leu Lys Gly Glu Val Ala Gln Met Ala
 275 280 285

Ala Tyr Tyr Glu His Arg Leu Gln Leu Gly Ser Lys Ala Ile Tyr His
 290 295 300

Leu Glu Ala Phe Val Ala Lys Phe Met Ala Leu Tyr Lys Lys Phe Ile
 305 310 315 320

Gln Asp Gly Leu Glu Gly Met Met Phe
 325

(2) ANGABEN ZU SEQ ID NO: 34:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 354 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

(A) ORGANISMUS: Homo sapiens

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 34:

Met Glu Val Glu Ala Val Cys Gly Gly Ala Gly Glu Val Glu Ala Gln
1 5 10 15

Asp Ser Asp Pro Ala Pro Ala Phe Ser Lys Ala Pro Gly Ser Ala Gly
 20 25 30

His Tyr Glu Leu Pro Trp Val Glu Lys Tyr Arg Pro Val Lys Leu Asn
 35 40 45

Glu Ile Val Gly Asn Glu Asp Thr Val Ser Arg Leu Glu Val Phe Ala
 50 55 60

Arg Glu Gly Asn Val Pro Asn Ile Ile Ile Ala Gly Pro Pro Gly Thr
65 70 75 80

Gly Lys Thr Thr Ser Ile Leu Cys Leu Ala Arg Ala Leu Leu Gly Pro
 85 90 95

Ala Leu Lys Asp Ala Met Leu Glu Leu Asn Ala Ser Asn Asp Arg Gly
 100 105 110

Ile Asp Val Val Arg Asn Lys Ile Lys Met Phe Ala Gln Gln Lys Val
 115 120 125

Thr Leu Pro Lys Gly Arg His Lys Ile Ile Ile Leu Asp Glu Ala Asp
 130 135 140

Ser Met Thr Asp Gly Ala Gln Gln Ala Leu Arg Arg Thr Met Glu Ile
145 150 155 160

Tyr Ser Lys Thr Thr Arg Phe Ala Leu Ala Cys Asn Ala Ser Asp Lys
 165 170 175

Ile Ile Glu Pro Ile Gln Ser Arg Cys Ala Val Leu Arg Tyr Thr Lys
 180 185 190

Leu Thr Asp Ala Gln Ile Leu Thr Arg Leu Met Asn Val Ile Glu Lys
 195 200 205

Glu Arg Val Pro Tyr Thr Asp Asp Gly Leu Glu Ala Ile Ile Phe Thr
 210 215 220

Ala Gln Gly Asp Met Arg Gln Ala Leu Asn Asn Leu Gln Ser Thr Phe
 225 230 235 240
 Ser Gly Phe Gly Phe Ile Asn Ser Glu Asn Val Phe Lys Val Cys Asp
 245 250 255
 Glu Pro His Pro Leu Leu Val Lys Glu Met Ile Gln His Cys Val Asn
 260 265 270
 Ala Asn Ile Asp Glu Ala Tyr Lys Ile Leu Ala His Leu Trp His Leu
 275 280 285
 Gly Tyr Ser Pro Glu Asp Ile Ile Gly Asn Ile Phe Arg Val Cys Lys
 290 295 300
 Thr Phe Gln Met Ala Glu Tyr Leu Lys Leu Glu Phe Ile Lys Glu Ile
 305 310 315 320
 Gly Tyr Thr His Met Lys Ile Ala Glu Gly Val Asn Ser Leu Leu Gln
 325 330 335
 Met Ala Gly Leu Leu Ala Arg Leu Cys Gln Lys Thr Met Ala Pro Val
 340 345 350
 Ala Ser

(2) ANGABEN ZU SEQ ID NO: 35:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 366 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(iii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: Escherichia coli

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 35:

Met Lys Phe Thr Val Glu Arg Glu His Leu Leu Lys Pro Leu Gln Gln
 1 5 10 15
 Val Ser Gly Pro Leu Gly Gly Arg Pro Thr Leu Pro Ile Leu Gly Asn
 20 25 30

Leu Leu Leu Gln Val Ala Asp Gly Thr Leu Ser Leu Thr Gly Thr Asp
 35 40 45

Leu Glu Met Glu Met Val Ala Arg Val Ala Leu Val Gln Pro His Glu
 50 55 60

Pro Gly Ala Thr Thr Val Pro Ala Arg Lys Phe Phe Asp Ile Cys Arg
 65 70 75 80

Gly Leu Pro Glu Gly Ala Glu Ile Ala Val Gln Leu Glu Gly Glu Arg
 85 90 95

Met Leu Val Arg Ser Gly Arg Ser Arg Phe Ser Leu Ser Thr Leu Pro
 100 105 110

Ala Ala Asp Phe Pro Asn Leu Asp Asp Trp Gln Ser Glu Val Glu Phe
 115 120 125

Thr Leu Pro Gln Ala Thr Met Lys Arg Leu Ile Glu Ala Thr Gln Phe
 130 135 140

Ser Met Ala His Gln Asp Val Arg Tyr Tyr Leu Asn Gly Met Leu Phe
 145 150 155 160

Glu Thr Glu Gly Glu Glu Leu Arg Thr Val Ala Thr Asp Gly His Arg
 165 170 175

Leu Ala Val Cys Ser Met Pro Ile Gly Gln Ser Leu Pro Ser His Ser
 180 185 190

Val Ile Val Pro Arg Lys Gly Val Ile Glu Leu Met Arg Met Leu Asp
 195 200 205

Gly Gly Asp Asn Pro Leu Arg Val Gln Ile Gly Ser Asn Asn Ile Arg
 210 215 220

Ala His Val Gly Asp Phe Ile Phe Thr Ser Lys Leu Val Asp Gly Arg
 225 230 235 240

Phe Pro Asp Tyr Arg Arg Val Leu Pro Lys Asn Pro Asp Lys His Leu
 245 250 255

Glu Ala Gly Cys Asp Leu Leu Lys Gln Ala Phe Ala Arg Ala Ala Ile
 260 265 270

Leu Ser Asn Glu Lys Phe Arg Gly Val Arg Leu Tyr Val Ser Glu Asn
 275 280 285

Gln Leu Lys Ile Thr Ala Asn Asn Pro Glu Gln Glu Glu Ala Glu Glu

290 295 300
 Ile Leu Asp Val Thr Tyr Ser Gly Ala Glu Met Glu Ile Gly Phe Asn
 305 310 315 320
 Val Ser Tyr Val Leu Asp Val Leu Asn Ala Leu Lys Cys Glu Asn Val
 325 330 335
 Arg Met Met Leu Thr Asp Ser Val Ser Ser Val Gln Ile Glu Asp Ala
 340 345 350
 Ala Ser Gln Ser Ala Ala Tyr Val Val Met Pro Met Arg Leu
 355 360 365

(2) ANGABEN ZU SEQ ID NO: 36:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 363 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: Aquifex Aeolicus

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 36:

Met Arg Val Lys Val Asp Arg Glu Glu Leu Glu Glu Val Leu Lys Lys
 1 5 10 15
 Ala Arg Glu Ser Thr Glu Lys Lys Ala Ala Leu Pro Ile Leu Ala Asn
 20 25 30
 Phe Leu Leu Ser Ala Lys Glu Glu Asn Leu Ile Val Arg Ala Thr Asp
 35 40 45
 Leu Glu Asn Tyr Leu Val Val Ser Val Lys Gly Glu Val Glu Glu Glu
 50 55 60
 Gly Glu Val Cys Val His Ser Gln Lys Leu Tyr Asp Ile Val Lys Asn
 65 70 75 80
 Leu Asn Ser Ala Tyr Val Tyr Leu His Thr Glu Gly Glu Lys Leu Val
 85 90 95

Ile Thr Gly Gly Lys Ser Thr Tyr Lys Leu Pro Thr Ala Pro Ala Glu
 100 105 110

Asp Phe Pro Glu Phe Pro Glu Ile Val Glu Gly Gly Glu Thr Leu Ser
 115 120 125

Gly Asn Leu Leu Val Asn Gly Ile Glu Lys Val Glu Tyr Ala Ile Ala
 130 135 140

Lys Glu Glu Ala Asn Ile Ala Leu Gln Gly Met Tyr Leu Arg Gly Tyr
 145 150 155 160

Glu Asp Arg Ile His Phe Val Gly Ser Asp Gly His Arg Leu Ala Leu
 165 170 175

Tyr Glu Pro Leu Gly Glu Phe Ser Lys Glu Leu Leu Ile Pro Arg Lys
 180 185 190

Ser Leu Lys Val Leu Lys Lys Leu Ile Thr Gly Ile Glu Asp Val Asn
 195 200 205

Ile Glu Lys Ser Glu Asp Glu Ser Phe Ala Tyr Phe Ser Thr Pro Glu
 210 215 220

Trp Lys Leu Ala Val Arg Leu Leu Glu Gly Glu Phe Pro Asp Tyr Met
 225 230 235 240

Ser Val Ile Pro Glu Glu Phe Ser Ala Glu Val Leu Phe Glu Thr Glu
 245 250 255

Glu Val Leu Lys Val Leu Lys Arg Leu Lys Ala Leu Ser Glu Gly Lys
 260 265 270

Val Phe Pro Val Lys Ile Thr Leu Ser Glu Asn Leu Ala Ile Phe Glu
 275 280 285

Phe Ala Asp Pro Glu Phe Gly Glu Ala Arg Glu Glu Ile Glu Val Glu
 290 295 300

Tyr Thr Gly Glu Pro Phe Glu Ile Gly Phe Asn Gly Lys Tyr Leu Met
 305 310 315 320

Glu Ala Leu Asp Ala Tyr Asp Ser Glu Arg Val Trp Phe Lys Phe Thr
 325 330 335

Thr Pro Asp Thr Ala Thr Leu Leu Glu Ala Glu Asp Tyr Glu Lys Glu
 340 345 350

Pro Tyr Lys Cys Ile Ile Met Pro Met Arg Val

355

360

(2) ANGABEN ZU SEQ ID NO: 37:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 1160 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: Escherichia coli

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 37:

Met Ser Glu Pro Arg Phe Val His Leu Arg Val His Ser Asp Tyr Ser
1 5 10 15

Met Ile Asp Gly Leu Ala Lys Thr Ala Pro Leu Val Lys Lys Ala Ala
20 25 30

Ala Leu Gly Met Pro Ala Leu Ala Ile Thr Asp Phe Thr Asn Leu Cys
35 40 45

Gly Leu Val Lys Phe Tyr Gly Ala Gly His Gly Ala Gly Ile Lys Pro
50 55 60

Ile Val Gly Ala Asp Phe Asn Val Gln Cys Asp Leu Leu Gly Asp Glu
65 70 75 80

Leu Thr His Leu Thr Val Leu Ala Ala Asn Asn Thr Gly Tyr Gln Asn
85 90 95

Leu Thr Leu Leu Ile Ser Lys Ala Tyr Gln Arg Gly Tyr Gly Ala Ala
100 105 110

Gly Pro Ile Ile Asp Arg Asp Trp Leu Ile Glu Leu Asn Glu Gly Leu
115 120 125

Ile Leu Leu Ser Gly Gly Arg Met Gly Asp Val Gly Arg Ser Leu Leu
130 135 140

Arg Gly Asn Ser Ala Leu Val Asp Glu Cys Val Ala Phe Tyr Glu Glu
145 150 155 160

His Phe Pro Asp Arg Tyr Phe Leu Glu Leu Ile Arg Thr Gly Arg Pro
 165 170 175

Asp Glu Glu Ser Tyr Leu His Ala Ala Val Glu Leu Ala Glu Ala Arg
 180 185 190

Gly Leu Pro Val Val Ala Thr Asn Asp Val Arg Phe Ile Asp Ser Ser
 195 200 205

Asp Phe Asp Ala His Glu Ile Arg Val Ala Ile His Asp Gly Phe Thr
 210 215 220

Leu Asp Asp Pro Lys Arg Pro Arg Asn Tyr Ser Pro Gln Gln Tyr Met
 225 230 235 240

Arg Ser Glu Glu Glu Met Cys Glu Leu Phe Ala Asp Ile Pro Glu Ala
 245 250 255

Leu Ala Asn Thr Val Glu Ile Ala Lys Arg Cys Asn Val Thr Val Arg
 260 265 270

Leu Gly Glu Tyr Phe Leu Pro Gln Phe Pro Thr Gly Asp Met Ser Thr
 275 280 285

Glu Asp Tyr Leu Val Lys Arg Ala Lys Glu Gly Leu Glu Glu Arg Leu
 290 295 300

Ala Phe Leu Phe Pro Asp Glu Glu Glu Arg Leu Lys Arg Arg Pro Glu
 305 310 315 320

Tyr Asp Glu Arg Leu Glu Thr Glu Leu Gln Val Ile Asn Gln Met Gly
 325 330 335

Phe Pro Gly Tyr Phe Leu Ile Val Met Glu Phe Ile Gln Trp Ser Lys
 340 345 350

Asp Asn Gly Val Pro Val Gly Pro Gly Arg Gly Ser Gly Ala Gly Ser
 355 360 365

Leu Val Ala Tyr Ala Leu Lys Ile Thr Asp Leu Asp Pro Leu Glu Phe
 370 375 380

Asp Leu Leu Phe Glu Arg Phe Leu Asn Pro Glu Arg Val Ser Met Pro
 385 390 395 400

Asp Phe Asp Val Asp Phe Cys Met Glu Lys Arg Asp Gln Val Ile Glu
 405 410 415

His Val Ala Asp Met Tyr Gly Arg Asp Ala Val Ser Gln Ile Ile Thr

420 425 430
 Phe Gly Thr Met Ala Ala Lys Ala Val Ile Arg Asp Val Gly Arg Val
 435 440 445
 Leu Gly His Pro Tyr Gly Phe Val Asp Arg Ile Ser Lys Leu Ile Pro
 450 455 460
 Pro Asp Pro Gly Met Thr Leu Ala Lys Ala Phe Glu Ala Glu Pro Gln
 465 470 475 480
 Leu Pro Glu Ile Tyr Glu Ala Asp Glu Glu Val Lys Ala Leu Ile Asp
 485 490 495
 Met Ala Arg Lys Leu Glu Gly Val Thr Arg Asn Ala Gly Lys His Ala
 500 505 510
 Gly Gly Val Val Ile Ala Pro Thr Lys Ile Thr Asp Phe Ala Pro Leu
 515 520 525
 Tyr Cys Asp Glu Glu Gly Lys His Pro Val Thr Gln Phe Asp Lys Ser
 530 535 540
 Asp Val Glu Tyr Ala Gly Leu Val Lys Phe Asp Phe Leu Gly Leu Arg
 545 550 555 560
 Thr Leu Thr Ile Ile Asn Trp Ala Leu Glu Met Ile Asn Lys Arg Arg
 565 570 575
 Ala Lys Asn Gly Glu Pro Pro Leu Asp Ile Ala Ala Ile Pro Leu Asp
 580 585 590
 Asp Lys Lys Ser Phe Asp Met Leu Gln Arg Ser Glu Thr Thr Ala Val
 595 600 605
 Phe Gln Leu Glu Ser Arg Gly Met Lys Asp Leu Ile Lys Arg Leu Gln
 610 615 620
 Pro Asp Cys Phe Glu Asp Met Ile Ala Leu Val Ala Leu Phe Arg Pro
 625 630 635 640
 Gly Pro Leu Gln Ser Gly Met Val Asp Asn Phe Ile Asp Arg Lys His
 645 650 655
 Gly Arg Glu Glu Ile Ser Tyr Pro Asp Val Gln Trp Gln His Glu Ser
 660 665 670
 Leu Lys Pro Val Leu Glu Pro Thr Tyr Gly Ile Ile Leu Tyr Gln Glu
 675 680 685

Gln Val Met Gln Ile Ala Gln Val Leu Ser Gly Tyr Thr Leu Gly Gly
690 695 700

Ala Asp Met Leu Arg Arg Ala Met Gly Lys Lys Lys Pro Glu Glu Met
705 710 715 720

Ala Lys Gln Arg Ser Val Phe Ala Glu Gly Ala Glu Lys Asn Gly Ile
725 730 735

Asn Ala Glu Leu Ala Met Lys Ile Phe Asp Leu Val Glu Lys Phe Ala
740 745 750

Gly Tyr Gly Phe Asn Lys Ser His Ser Ala Ala Tyr Ala Leu Val Ser
755 760 765

Tyr Gln Thr Leu Trp Leu Lys Ala His Tyr Pro Ala Glu Phe Met Ala
770 775 780

Ala Val Met Thr Ala Asp Met Asp Asn Thr Glu Lys Val Val Gly Leu
785 790 795 800

Val Asp Glu Cys Trp Arg Met Gly Leu Lys Ile Leu Pro Pro Asp Ile
805 810 815

Asn Ser Gly Leu Tyr His Phe His Val Asn Asp Asp Gly Glu Ile Val
820 825 830

Tyr Gly Ile Gly Ala Ile Lys Gly Val Gly Glu Gly Pro Ile Glu Ala
835 840 845

Ile Ile Glu Ala Arg Asn Lys Gly Gly Tyr Phe Arg Glu Leu Phe Asp
850 855 860

Leu Cys Ala Arg Thr Asp Thr Lys Lys Leu Asn Arg Arg Val Leu Glu
865 870 875 880

Lys Leu Ile Met Ser Gly Ala Phe Asp Arg Leu Gly Pro His Arg Ala
885 890 895

Ala Leu Met Asn Ser Leu Gly Asp Ala Leu Lys Ala Ala Asp Gln His
900 905 910

Ala Lys Ala Glu Ala Ile Gly Gln Ala Asp Met Phe Gly Val Leu Ala
915 920 925

Glu Glu Pro Glu Gln Ile Glu Gln Ser Tyr Ala Ser Cys Gln Pro Trp
930 935 940

Pro Glu Gln Val Val Leu Asp Gly Glu Arg Glu Thr Leu Gly Leu Tyr

945	950	955	960
Leu Thr Gly His Pro Ile Asn Gln Tyr Leu Lys Glu Ile Glu Arg Tyr			
965	970	975	
Val Gly Gly Val Arg Leu Lys Asp Met His Pro Thr Glu Arg Gly Lys			
980	985	990	
Val Ile Thr Ala Ala Gly Leu Val Val Ala Ala Arg Val Met Val Thr			
995	1000	1005	
Lys Arg Gly Asn Arg Ile Gly Ile Cys Thr Leu Asp Asp Arg Ser Gly			
1010	1015	1020	
Arg Leu Glu Val Met Leu Phe Thr Asp Ala Leu Asp Lys Tyr Gln Gln			
1025	1030	1035	1040
Leu Leu Glu Lys Asp Arg Ile Leu Ile Val Ser Gly Gln Val Ser Phe			
1045	1050	1055	
Asp Asp Phe Ser Gly Gly Leu Lys Met Thr Ala Arg Glu Val Met Asp			
1060	1065	1070	
Ile Asp Glu Ala Arg Glu Lys Tyr Ala Arg Gly Leu Ala Ile Ser Leu			
1075	1080	1085	
Thr Asp Arg Gln Ile Asp Asp Gln Leu Leu Asn Arg Leu Arg Gln Ser			
1090	1095	1100	
Leu Glu Pro His Arg Ser Gly Thr Ile Pro Val His Leu Tyr Tyr Gln			
1105	1110	1115	1120
Arg Ala Asp Ala Arg Ala Arg Leu Arg Phe Gly Ala Thr Trp Arg Val			
1125	1130	1135	
Ser Pro Ser Asp Arg Leu Leu Asn Asp Leu Arg Gly Leu Ile Gly Ser			
1140	1145	1150	
Glu Gln Val Glu Leu Glu Phe Asp			
1155	1160		

(2) ANGABEN ZU SEQ ID NO: 38:

(i) SEQUENZKENNZEICHEN:

(A) LÄNGE: 1161 Aminosäuren

(B) ART: Aminosäure

(C) STRANGFORM: Einzelstrang

(D) TOPOLOGIE: linear

(iii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNGLICHE HERKUNFT:

(A) ORGANISMUS: Aquifex Aeolicus

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 38:

Met Ser Lys Asp Phe Val His Leu His Leu His Thr Gln Phe Ser Leu
1 5 10 15

Leu Asp Gly Ala Ile Lys Ile Asp Glu Leu Val Lys Lys Ala Lys Glu
20 25 30

Tyr Gly Tyr Lys Ala Val Gly Met Ser Asp His Gly Asn Leu Phe Gly
35 40 45

Ser Tyr Lys Phe Tyr Lys Ala Leu Lys Ala Glu Gly Ile Lys Pro Ile
50 55 60

Ile Gly Met Glu Ala Tyr Phe Thr Thr Gly Ser Arg Phe Asp Arg Lys
65 70 75 80

Thr Lys Thr Ser Glu Asp Asn Ile Thr Asp Lys Tyr Asn His His Leu
85 90 95

Ile Leu Ile Ala Lys Asp Asp Lys Gly Leu Lys Asn Leu Met Lys Leu
100 105 110

Ser Thr Leu Ala Tyr Lys Glu Gly Phe Tyr Tyr Lys Pro Arg Ile Asp
115 120 125

Tyr Glu Leu Leu Glu Lys Tyr Gly Glu Gly Leu Ile Ala Leu Thr Ala
130 135 140

Cys Leu Lys Gly Val Pro Thr Tyr Tyr Ala Ser Ile Asn Glu Val Lys
145 150 155 160

Lys Ala Glu Glu Trp Val Lys Lys Phe Lys Asp Ile Phe Gly Asp Asp
165 170 175

Leu Tyr Leu Glu Leu Gln Ala Asn Asn Ile Pro Glu Gln Glu Val Ala
180 185 190

Asn Arg Asn Leu Ile Glu Ile Ala Lys Lys Tyr Asp Val Lys Leu Ile
195 200 205

Ala Thr Gln Asp Ala His Tyr Leu Asn Pro Glu Asp Arg Tyr Ala His

210	215	220	
Thr Val Leu Met Ala Leu Gln Met Lys Lys Thr Ile His Glu Leu Ser			
225	230	235	240
Ser Gly Asn Phe Lys Cys Ser Asn Glu Asp Leu His Phe Ala Pro Pro			
	245	250	255
Glu Tyr Met Trp Lys Lys Phe Glu Gly Lys Phe Glu Gly Trp Glu Lys			
	260	265	270
Ala Leu Leu Asn Thr Leu Glu Val Met Glu Lys Thr Ala Asp Ser Phe			
	275	280	285
Glu Ile Phe Glu Asn Ser Thr Tyr Leu Leu Pro Lys Tyr Asp Val Pro			
	290	295	300
Pro Asp Lys Thr Leu Glu Glu Tyr Leu Arg Glu Leu Ala Tyr Lys Gly			
	305	310	315
Leu Arg Gln Arg Ile Glu Arg Gly Gln Ala Lys Asp Thr Lys Glu Tyr			
	325	330	335
Trp Glu Arg Leu Glu Tyr Glu Leu Glu Val Ile Asn Lys Met Gly Phe			
	340	345	350
Ala Gly Tyr Phe Leu Ile Val Gln Asp Phe Ile Asn Trp Ala Lys Lys			
	355	360	365
Asn Asp Ile Pro Val Gly Pro Gly Arg Gly Ser Ala Gly Gly Ser Leu			
	370	375	380
Val Ala Tyr Ala Ile Gly Ile Thr Asp Val Asp Pro Ile Lys His Gly			
	385	390	395
Phe Leu Phe Glu Arg Phe Leu Asn Pro Glu Arg Val Ser Met Pro Asp			
	405	410	415
Ile Asp Val Asp Phe Cys Gln Asp Asn Arg Glu Lys Val Ile Glu Tyr			
	420	425	430
Val Arg Asn Lys Tyr Gly His Asp Asn Val Ala Gln Ile Ile Thr Tyr			
	435	440	445
Asn Val Met Lys Ala Lys Gln Thr Leu Arg Asp Val Ala Arg Ala Met			
	450	455	460
Gly Leu Pro Tyr Ser Thr Ala Asp Lys Leu Ala Lys Leu Ile Pro Gln			
	465	470	475
			480

Gly Asp Val Gln Gly Thr Trp Leu Ser Leu Glu Glu Met Tyr Lys Thr
 485 490 495

Pro Val Glu Glu Leu Leu Gln Lys Tyr Gly Glu His Arg Thr Asp Ile
 500 505 510

Glu Asp Asn Val Lys Lys Phe Arg Gln Ile Cys Glu Glu Ser Pro Glu
 515 520 525

Ile Lys Gln Leu Val Glu Thr Ala Leu Lys Leu Glu Gly Leu Thr Arg
 530 535 540

His Thr Ser Leu His Ala Ala Gly Val Val Ile Ala Pro Lys Pro Leu
 545 550 555 560

Ser Glu Leu Val Pro Leu Tyr Tyr Asp Lys Glu Gly Glu Val Ala Thr
 565 570 575

Gln Tyr Asp Met Val Gln Leu Glu Glu Leu Gly Leu Leu Lys Met Asp
 580 585 590

Phe Leu Gly Leu Lys Thr Leu Thr Glu Leu Lys Leu Met Lys Glu Leu
 595 600 605

Ile Lys Glu Arg His Gly Val Asp Ile Asn Phe Leu Glu Leu Pro Leu
 610 615 620

Asp Asp Pro Lys Val Tyr Lys Leu Leu Gln Glu Gly Lys Thr Thr Gly
 625 630 635 640

Val Phe Gln Leu Glu Ser Arg Gly Met Lys Glu Leu Leu Lys Lys Leu
 645 650 655

Lys Pro Asp Ser Phe Asp Asp Ile Val Ala Val Leu Ala Leu Tyr Arg
 660 665 670

Pro Gly Pro Leu Lys Ser Gly Leu Val Asp Thr Tyr Ile Lys Arg Lys
 675 680 685

His Gly Lys Glu Pro Val Glu Tyr Pro Phe Pro Glu Leu Glu Pro Val
 690 695 700

Leu Lys Glu Thr Tyr Gly Val Ile Val Tyr Gln Glu Gln Val Met Lys
 705 710 715 720

Met Ser Gln Ile Leu Ser Gly Phe Thr Pro Gly Glu Ala Asp Thr Leu
 725 730 735

Arg Lys Ala Ile Gly Lys Lys Lys Ala Asp Leu Met Ala Gln Met Lys

740	745	750
Asp Lys Phe Ile Gln Gly Ala Val Glu Arg Gly Tyr Pro Glu Glu Lys		
755	760	765
Ile Arg Lys Leu Trp Glu Asp Ile Glu Lys Phe Ala Ser Tyr Ser Phe		
770	775	780
Asn Lys Ser His Ser Val Ala Tyr Gly Tyr Ile Ser Tyr Trp Thr Ala		
785	790	795 800
Tyr Val Lys Ala His Tyr Pro Ala Glu Phe Phe Ala Val Lys Leu Thr		
805	810	815
Thr Glu Lys Asn Asp Asn Lys Phe Leu Asn Leu Ile Lys Asp Ala Lys		
820	825	830
Leu Phe Gly Phe Glu Ile Leu Pro Pro Asp Ile Asn Lys Ser Asp Val		
835	840	845
Gly Phe Thr Ile Glu Gly Glu Asn Arg Ile Arg Phe Gly Leu Ala Arg		
850	855	860
Ile Lys Gly Val Gly Glu Glu Thr Ala Lys Ile Ile Val Glu Ala Arg		
865	870	875 880
Lys Lys Tyr Lys Gln Phe Lys Gly Leu Ala Asp Phe Ile Asn Lys Thr		
885	890	895
Lys Asn Arg Lys Ile Asn Lys Lys Val Val Glu Ala Leu Val Lys Ala		
900	905	910
Gly Ala Phe Asp Phe Thr Lys Lys Lys Arg Lys Glu Leu Leu Ala Lys		
915	920	925
Val Ala Asn Ser Glu Lys Ala Leu Met Ala Thr Gln Asn Ser Leu Phe		
930	935	940
Gly Ala Pro Lys Glu Glu Val Glu Glu Leu Asp Pro Leu Lys Leu Glu		
945	950	955 960
Lys Glu Val Leu Gly Phe Tyr Ile Ser Gly His Pro Leu Asp Asn Tyr		
965	970	975
Glu Lys Leu Leu Lys Asn Arg Tyr Thr Pro Ile Glu Asp Leu Glu Glu		
980	985	990
Trp Asp Lys Glu Ser Glu Ala Val Leu Thr Gly Val Ile Thr Glu Leu		
995	1000	1005

Lys Val Lys Lys Thr Lys Asn Gly Asp Tyr Met Ala Val Phe Asn Leu
1010 1015 1020

Val Asp Lys Thr Gly Leu Ile Glu Cys Val Val Phe Pro Gly Val Tyr
1025 1030 1035 1040

Glu Glu Ala Lys Glu Leu Ile Glu Glu Asp Arg Val Val Val Val Lys
1045 1050 1055

Gly Phe Leu Asp Glu Asp Leu Glu Thr Glu Asn Val Lys Phe Val Val
1060 1065 1070

Lys Glu Val Phe Ser Pro Glu Glu Phe Ala Lys Glu Met Arg Asn Thr
1075 1080 1085

Leu Tyr Ile Phe Leu Lys Arg Glu Gln Ala Leu Asn Gly Val Ala Glu
1090 1095 1100

Lys Leu Lys Gly Ile Ile Glu Asn Asn Arg Thr Glu Asp Gly Tyr Asn
1105 1110 1115 1120

Leu Val Leu Thr Val Asp Leu Gly Asp Tyr Phe Val Asp Leu Ala Leu
1125 1130 1135

Pro Gln Asp Met Lys Leu Lys Ala Asp Arg Lys Val Val Glu Glu Ile
1140 1145 1150

Glu Lys Leu Gly Val Lys Val Ile Ile
1155 1160

(2) ANGABEN ZU SEQ ID NO: 39:

- (i) SEQUENZKENNZEICHEN:
 - (A) LÄNGE: 64 Aminosäuren
 - (B) ART: Aminosäure
 - (C) STRANGFORM: Einzelstrang
 - (D) TOPOLOGIE: beides

- (ii) ART DES MOLEKÜLS: Protein

- (vi) URSPRÜNGLICHE HERKUNFT:
 - (A) ORGANISMUS: Synthetisch

xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 39:

[GAVLIMPFW]-D-X-X-X-[GAVLIMPFW]-X-X-[GAVLIMPFW]-X-[GAVLIMPFW]-X-
[GAVLIMPFW]-X-X-X-X-F-X-X-Y-X-X-D

64

(2) ANGABEN ZU SEQ ID NO: 40:

- (i) SEQUENZKENNZEICHEN:
 - (A) LÄNGE: 26 Aminosäuren
 - (B) ART: Aminosäure
 - (C) STRANGFORM: Einzelstrang
 - (D) TOPOLOGIE: beides

- (ii) ART DES MOLEKÜLS: Protein

- (vi) URSPRÜNGLICHE HERKUNFT:
 - (A) ORGANISMUS: Synthetisch

xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 40:

[GAVLIMPFW]-X(3)-L-A-P-[KRHDE]-[GAVLIMPFW]-E

28

(2) ANGABEN ZU SEQ ID NO: 41:

- (i) SEQUENZKENNZEICHEN:
 - (A) LÄNGE: 51 Aminosäuren
 - (B) ART: Aminosäure
 - (C) STRANGFORM: Einzelstrang
 - (D) TOPOLOGIE: beides

- (ii) ART DES MOLEKÜLS: Protein

- (vi) URSPRÜNGLICHE HERKUNFT:
 - (A) ORGANISMUS: Synthetisch

xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 41:

C-N-Y-X-S-[KRHDE]-I-I-X-[GAVLIMPFW]-[GAVLIMPFW]-Q-S-R-C-X-X-F-R-F-X-P-

[GAVLIMPFW]

51

(2) ANGABEN ZU SEQ ID NO: 42:

- (i) SEQUENZKENNZEICHEN:
 - (A) LÄNGE: 41 Aminosäuren
 - (B) ART: Aminosäure
 - (C) STRANGFORM: Einzelstrang
 - (D) TOPOLOGIE: beides

(ii) ART DES MOLEKÜLS: Protein

- (vi) URSPRÜNGLICHE HERKUNFT:
 - (A) ORGANISMUS: Synthetisch

xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 42:

K-X-X-L-L-X-G-P-P-G-X-G-X-T-[STNQYC]-X-[GAVLIMPFW]-X-X-[GAVLIMPFW] 41

(2) ANGABEN ZU SEQ ID NO: 43:

- (i) SEQUENZKENNZEICHEN:
 - (A) LÄNGE: 80 Aminosäuren
 - (B) ART: Aminosäure
 - (C) STRANGFORM: Einzelstrang
 - (D) TOPOLOGIE: beides

(ii) ART DES MOLEKÜLS: Protein

- (vi) URSPRÜNGLICHE HERKUNFT:
 - (A) ORGANISMUS: Synthetisch

xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 43:

[FL]-[GAVLIMPFW]-X-X-[GAVLIMPFW]-X-G-X(13)-[GAVLIMPFW]-X-[YR]-
[GAVLIMPFW]-X-[GAVLIMPFW]-A-G-[DN]-[GAVLIMPFW]-[GAVLIMPFW]-[DS] 80

(2) ANGABEN ZU SEQ ID NO: 44:

- (i) SEQUENZKENNZEICHEN:
 - (A) LÄNGE: 44 Aminosäuren
 - (B) ART: Aminosäure
 - (C) STRANGFORM: Einzelstrang
 - (D) TOPOLOGIE: beides

(ii) ART DES MOLEKÜLS: Protein

- (vi) URSPRÜNGLICHE HERKUNFT:
 - (A) ORGANISMUS: Synthetisch

xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 44:

D-[GAVLIMPFW]-[GAVLIMPFW]-X-X-Y-N-X-X-X-F-D-X-P-Y-[GAVLIMPFW]-X-X-R-A 44

(2) ANGABEN ZU SEQ ID NO: 45:

- (i) SEQUENZKENNZEICHEN:
(A) LÄNGE: 78 Aminosäuren
(B) ART: Aminosäure
(C) STRANGFORM: Einzelstrang
(D) TOPOLOGIE: beides

(ii) ART DES MOLEKÜLS: Protein

- (vi) URSPRÜNGLICHE HERKUNFT:
(A) ORGANISMUS: Synthetisch

xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 45:

A-[GAVLIMPFW]-R-T-A-[GAVLIMPFW]-A-[GAVLIMPFW]-[GAVLIMPFW]-T-E-G-
[GAVLIMPFW]-V-X-A-P-[GAVLIMPFW]-E-G-I-A-X-V-[KRHDE]-I

78

(2) ANGABEN ZU SEQ ID NO: 46:

- (i) SEQUENZKENNZEICHEN:
(A) LÄNGE: 118 Aminosäuren
(B) ART: Aminosäure
(C) STRANGFORM: Einzelstrang
(D) TOPOLOGIE: beides

(ii) ART DES MOLEKÜLS: Protein

- (vi) URSPRÜNGLICHE HERKUNFT:
(A) ORGANISMUS: Synthetisch

xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 46:

[GAVLIMPFW]-P-V-G-[GAVLIMPFW]-G-R-G-S-X-[GAVLIMPFW]-G-S-[GAVLIMPFW]-V-A-
X-A-[GAVLIMPFW]-X-I-T-D-[GAVLIMPFW]-D-P-[GAVLIMPFW]-X-X-X-[GAVLIMPFW]-L-
F-E-R-F-L-N-P-E-R-[GAVLIMPFW]-S-M-P-D

118

(2) ANGABEN ZU SEQ ID NO: 47:

- (i) SEQUENZKENNZEICHEN:
(A) LÄNGE: 29 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Einzelstrang
(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: DNA

- (vi) URSPRÜNGLICHE HERKUNFT:
(A) ORGANISMUS: M13 MP18 ss DNA (Phage)

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 47:

GCATTGACCG TAATGGGATA GGTACGTT

29

(2) ANGABEN ZU SEQ ID NO: 46:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 29 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: DNA

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: M13 MP18 ss DNA (Phage)

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 48:

AGCGGATAAC AATTCACAC AGGAAACAG

29

(2) ANGABEN ZU SEQ ID NO: 49:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 32 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: DNA

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: Archaeoglobus fulgidus

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 49:

ACGCCGGGAT CCATAGACGT CATAATGACC GG

32

(2) ANGABEN ZU SEQ ID NO: 50:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 31 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: DNA

- (vi) URSPRÜNGLICHE HERKUNFT:
(A) ORGANISMUS: *Archaeoglobus fulgidus*

- (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 50:

TACGGGGTAC CCGAGCCAAA ATTGGGTAAA G

31

- (2) ANGABEN ZU SEQ ID NO: 51:

- (i) SEQUENZKENNZEICHEN:
(A) LÄNGE: 32 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Einzelstrang
(D) TOPOLOGIE: linear

- (ii) ART DES MOLEKÜLS: DNA

- (vi) URSPRÜNGLICHE HERKUNFT:
(A) ORGANISMUS: *Archaeoglobus fulgidus*

- (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 51:

ACGCGCGGAT CCATAGACGT CATAATGACC GG

32

- (2) ANGABEN ZU SEQ ID NO: 52:

- (i) SEQUENZKENNZEICHEN:
(A) LÄNGE: 31 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Einzelstrang
(D) TOPOLOGIE: linear

- (ii) ART DES MOLEKÜLS: DNA

- (vi) URSPRÜNGLICHE HERKUNFT:
(A) ORGANISMUS: *Archaeoglobus fulgidus*

- (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 52:

TACGGGGTAC CCGAGCCAAA ATTGGGTAAA G

31

- (2) ANGABEN ZU SEQ ID NO: 53:

- (i) SEQUENZKENNZEICHEN:
(A) LÄNGE: 19 Basenpaar

- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear
- (ii) ART DES MOLEKÜLS: DNA
- (vi) URSPRÜNGLICHE HERKUNFT:
 - (A) ORGANISMUS: Mensch Kollagen Forward
- (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 53:
TAAAGGGTCA CCGTGGTTC

19

(2) ANGABEN ZU SEQ ID NO: 54:

- (i) SEQUENZKENNZEICHEN:
 - (A) LÄNGE: 20 Basenpaare
 - (B) ART: Nukleinsäure
 - (C) STRANGFORM: Einzelstrang
 - (D) TOPOLOGIE: linear
- (ii) ART DES MOLEKÜLS: DNA
- (vi) URSPRÜNGLICHE HERKUNFT:
 - (A) ORGANISMUS: Mensch Kollagen Reverse

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 54:

CGAACCACAT TGGCATCATC

20